HARVARD BUSINESS REVIEW PRESS

Đào Lợi *dịch*

Ajay Agrawal | Joshua Gans | Avi Goldfarb

AI TRONG CUỘC CÁCH MẠNG CÔNG NGHỆ 4.0

Con đường ngắn nhất để phát triển doanh nghiệp



NHÀ XUẤT BẢN LAO ĐỘNG

Mục lục

Tới gia đình của chúng tôi, những người đồng nghiệp, những người học trò và những công ty khởi nghiệp đã truyền cảm hứng để chúng tôi suy nghĩ sáng suốt và sâu sắc về trí tuệ nhân tạo.

1Giới thiệu - Trí tuệ nhân tạo

N

ếu viễn cảnh sau đây nghe chưa quen thuộc, thì nó sẽ sớm thôi. Một đứa trẻ làm bài tập về nhà một mình ở trong phòng. Một câu hỏi vang lên: “Đâu là thủ phủ của Delaware?”. Bố mẹ đứa trẻ bắt đầu nghĩ: “Baltimore… rõ ràng là vậy mà… Wilmington… không phải là thủ phủ.” Nhưng trước khi bố mẹ đứa trẻ kịp nghĩ ra, một cỗ máy thông minh tên Alexa đã nói ra câu trả lời chính xác: “Thủ phủ của Delaware là Dover.” Alexa là chiếc máy trí tuệ nhân tạo của Amazon, hay còn gọi là AI, nó có thể phiên dịch ngôn ngữ tự nhiên và cung cấp câu trả lời cho những câu hỏi với tốc độ ánh sáng. Alexa đã thay thế các bậc phụ huynh với tư cách là nguồn thông tin dồi dào dưới con mắt của con trẻ.

AI có mặt ở mọi nơi. AI ở trong điện thoại, xe ô tô của chúng ta, ở trong những trải nghiệm mua sắm, hẹn hò, ở trong bệnh viện, ngân hàng và ở trên các tin tức. Vì vậy, chẳng lạ gì khi các giám đốc doanh nghiệp, giám đốc điều hành, phó chủ tịch, quản lý, trưởng nhóm, các nhà khởi nghiệp, nhà đầu tư, tư vấn viên, nhà hoạch định chính sách đang trong cuộc chạy đua để tìm hiểu về AI: họ đều nhận ra rằng nó sẽ gây ra những thay đổi thiết yếu lên doanh nghiệp của họ.

Ba người chúng tôi đã quan sát những sự phát triển của AI từ một điểm nhìn đặc biệt có lợi. Chúng tôi là những nhà kinh tế học đã xây dựng sự nghiệp nhờ việc nghiên cứu cách mạng công nghệ tuyệt vời nhất: Internet. Qua nhiều năm nghiên cứu, chúng tôi đã học cách cắt lược những sự cường điệu hoá để tập trung vào ý nghĩa của công nghệ với những người đưa ra quyết định.

Chúng tôi cũng đã xây dựng Creative Destruction Lab (CDL), một chương trình đang ở giai đoạn khởi đầu nhằm giúp tăng khả năng thành công cho các dự án khởi nghiệp liên quan tới khoa học. Ban đầu, CDL được dành cho tất cả các dự án khởi nghiệp, nhưng tới năm 2015, rất nhiều trong số những thương vụ đáng mong chờ nhất đều là những công ty được vận hành nhờ AI. Cho tới tháng 9 năm 2017, CDL đã trở thành một công ty tập trung nhiều nhất các dự án khởi nghiệp về AI trên thế giới trong vòng ba năm liên tiếp.

Bởi vậy, rất nhiều lãnh đạo trong ngành thường xuyên tới Toronto để tham gia vào CDL. Ví dụ, một trong những nhà đầu tư chính của chương trình AI nền tảng cho Alexa Amazon, William Tunstall-Pedoe, cứ cách tám tuần lại bay từ Cambridge, Anh tới Toronto để tham gia cùng chúng tôi trong suốt thời gian dự án. Barney Pell ở San Francisco, người đã từng đứng đầu một đội gồm 85 người ở NASA phóng AI đầu tiên vào vũ trụ, cũng vậy.

Lí do cho ưu thế của CDL trong lĩnh vực này một phần do địa điểm của chúng tôi ở Toronto, nơi mà rất nhiều phát minh nòng cốt mang lại nhiều sự quan tâm gần đây tới AI – trong lĩnh vực được gọi là “máy tự học” – được ươm mầm và nuôi dưỡng. Những chuyên gia đã từng làm việc tại bộ phận khoa học máy tính tại Đại học Toronto hiện đang dẫn dắt một số đội ngũ công nghiệp AI hàng đầu trên thế giới, bao gồm những nhân tài ở Facebook, Apple và Open AI của Elon Musk.

Vì được tiếp xúc với rất nhiều *ứng dụng* của AI, chúng tôi có động lực tập trung nghiên cứu về sự ảnh hưởng đến chiến lược kinh doanh của loại hình công nghệ này. Như chúng tôi sẽ lí giải ở các chương sau, AI là công nghệ mang tính dự đoán, những dự đoán là thông tin đầu vào của quá trình quyết định và nền kinh tế đã cung cấp một khuôn mẫu hoàn hảo để hiểu rõ hơn về những sự đánh đổi nằm sau mỗi quyết định. Vậy nên, nhờ sự may mắn, chúng tôi nhận thấy bản thân đang ở đúng chỗ, đúng thời điểm để tạo cầu nối giữa chuyên gia công nghệ và người làm kinh doanh. Cuốn sách này chính là thành quả của chúng tôi.

Sự nhìn nhận quan trọng đầu tiên của chúng tôi là làn sóng mới của trí tuệ nhân tạo không thực sự mang lại cho chúng ta trí tuệ mà là một thành phần quan trọng của trí tuệ - *sự dự đoán.* Điều mà Alexa đã làm khi một đứa trẻ đặt ra một câu hỏi là nghe thấy âm thanh, dự đoán những từ ngữ mà đứa trẻ nói rồi dự đoán thông tin mà những từ ngữ đang tìm kiếm. Alexa không “biết” thủ phủ của Delaware là gì. Nhưng Alexa có khả năng dự đoán, khi con người hỏi một câu hỏi như vậy, nó biết tìm kiếm câu trả lời cụ thể: “Dover”.

Mỗi dự án khởi nghiệp trong phòng thí nghiệm của chúng tôi đều dựa vào việc khai thác những lợi ích của việc dự đoán tốt hơn. Deep Genomics nâng cao hoạt tính của thuốc thông qua việc dự đoán điều gì sẽ xảy ra trong một tế bào khi ADN bị biến đổi. Chisel nâng cao tính năng của pháp luật thông qua việc dự đoán phần nào trong văn bản cần biên soạn lại. Validere nâng cao

hiệu quả của chuyển giao dầu bằng cách dự đoán hàm lượng nước của dầu thô. Những ứng dụng này chỉ là một phần nhỏ so với hầu hết những điều mà các doanh nghiệp sẽ làm trong tương lai gần.

Nếu bạn bị đang lạc lối trong việc tìm ra ý nghĩa của AI, chúng tôi có thể giúp bạn hiểu được ý nghĩa của AI và dẫn lối bạn tìm hiểu những tiến bộ trong công nghệ này.

Nếu bạn là một nhà lãnh đạo doanh nghiệp, chúng tôi sẽ giúp bạn hiểu về tác động của AI đối với công tác quản lý và các quyết định. Nếu bạn là sinh viên hay vừa mới tốt nghiệp, chúng tôi sẽ cung cấp cho bạn một khuôn mẫu để bạn suy nghĩ về sự tiến hóa của các công việc và ngành nghề trong tương lai. Nếu bạn là một chuyên gia phân tích tài chính hoặc nhà đầu tư mạo hiểm, chúng tôi sẽ cung cấp một cấu trúc mà dựa vào đó bạn có thể phát triển các luận điểm đầu tư.

Nếu bạn là nhà hoạch định chính sách, chúng tôi sẽ cung cấp hướng dẫn để bạn hiểu cách AI có thể thay đổi xã hội ra sao và những chính sách có thể định hình những thay đổi mang hướng tốt hơn như thế nào.

Kinh tế cung cấp một nền tảng vững chắc để hiểu về sự bất định và ý nghĩa của nó đối với quá trình đưa ra quyết định. Vì dự đoán tốt hơn sẽ làm giảm sự do dự, chúng tôi sử dụng nền kinh tế để giúp bạn hiểu rõ ý nghĩa của AI đối với các quyết định bạn đưa ra trong quá trình kinh doanh của bạn. Điều này sẽ cung cấp sự nhận thức sâu sắc về những công cụ AI nào có khả năng mang lại lợi nhuận đầu tư cao nhất cho dòng chảy trong công việc kinh doanh của bạn. Từ đó dẫn đến một khuôn mẫu cho việc xây dựng chiến lược kinh doanh, chẳng hạn như cách bạn có thể suy nghĩ lại quy mô và phạm vi doanh nghiệp của mình để khai thác những vấn đề thực tế của kinh tế mới dựa trên dự đoán. Cuối cùng, chúng tôi chia nhỏ những sự đánh đổi liên quan đến AI trong công việc, trong sự tập trung quyền lực của sức mạnh đoàn thể, trong quyền riêng tư và trong địa chính trị.

Những dự đoán nào quan trọng đối với doanh nghiệp của bạn? Những tiến bộ trong AI sẽ thay đổi những dự đoán của bạn như thế nào? Ngành của bạn sẽ thiết kế lại các công việc như thế nào để đáp ứng những tiến bộ trong công nghệ dự đoán, giống như cách các ngành nghề định hình lại công việc với sự phát triển của máy tính cá nhân và rồi là Internet? AI vẫn còn là một khái

niệm mới và vẫn chưa được hiểu một cách rõ ràng, nhưng công cụ kinh tế để đánh giá tác động của việc giảm chi phí dự đoán là chắc chắn; mặc dù những ví dụ chúng tôi sử dụng rồi cũng sẽ trở nên lỗi thời, nhưng khuôn mẫu được đề cập trong cuốn sách này thì không. Những sự nhận thức sâu sắc sẽ tiếp tục được áp dụng khi công nghệ phát triển và những dự đoán trở nên chính xác, phức tạp hơn.

Trí tuệ nhân tạo và cuộc cách mạng trong kinh doanh không phải là công thức để thành công trong nền kinh tế AI. Thay vào đó, chúng tôi nhấn mạnh tới những sự đánh đổi liên quan đến AI. Nhiều dữ liệu hơn đồng nghĩa với việc ít quyền riêng tư hơn. Tốc độ nhanh hơn đồng nghĩa với độ chính xác thấp hơn. Quyền tự chủ cao hơn đồng nghĩa với việc ít sự kiểm soát hơn. Chúng tôi không áp đặt chiến lược tốt nhất cho doanh nghiệp của bạn. Đó là công việc của bạn. Chiến lược tốt nhất cho doanh nghiệp, nghề nghiệp hoặc đất nước của bạn sẽ dựa vào cách bạn cân nhắc mỗi mặt của sự đánh đổi. Cuốn sách này cung cấp cho bạn khuôn mẫu để xác định những sự đánh đổi chính và cách đánh giá ưu điểm, nhược điểm để bạn có thể đi tới quyết định tốt nhất. Tất nhiên, ngay cả khi đã có khuôn mẫu của chúng tôi trong tay, bạn vẫn sẽ thấy mọi thứ thay đổi nhanh chóng. Bạn sẽ phải đưa ra quyết định ngay cả khi không có thông tin đầy đủ, nhưng làm vậy còn tốt hơn là không làm gì.

NHỮNG ĐIỂM CHÍNH

• Làn sóng mới của trí tuệ nhân tạo không thực sự mang đến cho chúng ta trí tuệ mà thay vào đó là một thành phần quan trọng của trí tuệ - sự dự đoán

• Sự dự đoán là dữ liệu đầu vào quan trọng của quá trình đưa ra quyết định. Nền kinh tế cho chúng ta một khuôn mẫu đã được xây dựng đầy đủ để hiểu rõ về quá trình đưa ra quyết định. Những tác động còn mới và chưa được hiểu rõ trong sự tiến bộ của công nghệ dự đoán có thể được kết hợp cùng với logic lý thuyết quyết định cũ và quen thuộc của nền kinh tế để mang lại những sự thấu hiểu sâu sắc, giúp định hướng cách tiếp cận AI cho tổ chức của bạn.

• Thường sẽ không có một câu trả lời đúng cho câu hỏi đâu là kế hoạch AI tốt nhất hay những công cụ AI nào tốt nhất, vì AI bao gồm những sự đánh đổi: tốc độ nhanh hơn thì ít chính xác hơn; quyền tự chủ cao hơn thì ít sự kiểm soát hơn; nhiều dữ liệu hơn thì ít quyền riêng tư hơn. Chúng tôi cung cấp cho

bạn một phương pháp để xác định những sự đánh đổi với mỗi quyết định liên quan tới AI để bạn có thể đánh giá hai mặt ưu – nhược của mỗi sự đánh đổi sao cho phù hợp với sứ mệnh và mục tiêu của tổ chức, từ đó đưa ra quyết định tốt nhất cho bạn.

2Giá thành rẻ thay đổi tất cả

A

i cũng đã từng có hoặc sẽ sớm tiếp xúc với AI. Chúng ta đã quen với phương tiện truyền thông ngập tràn những câu chuyện về công nghệ mới sẽ thay đổi cuộc sống của chúng ta. Trong khi những người là dân kỹ thuật vui mừng trước những tiềm năng mới trong tương lai và những người bài xích công nghệ thì thương nhớ những ngày tốt đẹp đã qua, đa số chúng ta đã quen với nhịp thay đổi của các thông tin về công nghệ đến mức chúng ta vô cảm mà nói rằng điều duy nhất miễn nhiễm với sự thay đổi chính là thay đổi. Cho tới khi chúng ta tiếp xúc với AI. Lúc đó chúng ta sẽ nhận ra rằng công nghệ này thực sự khác biệt.

Một vài nhà khoa học máy tính có trải nghiệm tiếp xúc với AI vào năm 2012 khi một nhóm học sinh tới từ Đại học Toronto chiến thắng đầy ấn tượng trong cuộc thi nhận dạng đồ vật bằng thị giác mang tên Image-Net, họ khiến tất cả những đội lọt vào chung kết năm sau sử dụng cách tiếp cận học sâu - được coi là mới ở thời điểm đó - để cạnh tranh. Nhận dạng đồ vật không chỉ đơn thuần là một trò chơi; công nghệ đó giúp các loại máy móc có thể "nhìn" được.

Một số giám đốc điều hành công nghệ tiếp xúc với AI khi họ đọc được tiêu đề báo vào tháng 1 năm 2014 về việc Google vừa mới trả hơn 600 triệu đô la để thu mua công ty DeepMind tại Anh, mặc dù công ty khởi nghiệp mới này chỉ tạo ra doanh thu không đáng kể so với giá mua nhưng đã thể hiện được rằng, công nghệ AI của họ đã tự học, mà không cần phải lập trình, để chơi một số trò chơi của Atari với hiệu suất nhanh hơn con người.

Một vài công dân bình thường đã có trải nghiệm tiếp xúc với AI sau khi nhà vật lý học nổi tiếng Stephen Hawking giải thích nhấn mạnh rằng: "Tất cả những điều mà nhân loại có thể đem lại chính là sản phẩm của trí tuệ con người… Thành công trong việc tạo ra AI sẽ là sự kiện lớn nhất trong lịch sử loài người."[1]

Một số khác đã có trải nghiệm tiếp xúc với AI lần đầu khi họ không cần dùng

tay để điều khiển chiếc Tesla đang chạy và điều hướng giao thông với công nghệ lái xe tự động (AI Autopilot).

Chính phủ Trung Quốc đã tiếp xúc với AI khi họ chứng kiến công nghệ AI của DeepMind, AlphaGo, đánh bại Lee Se-dol, kì thủ cờ vây bậc thầy người Hàn Quốc , rồi sau đó cùng năm đánh bại kì thủ giỏi hàng đầu thế giới, Ke Jie của Trung Quốc. Báo *New York Times* gọi trận đấu đó như là "thời khắc Sputnik"[2]của Trung Quốc. Giống như sự đầu tư hàng loạt của Mỹ vào khoa học sau sự kiện Liên Xô ra mắt tên lửa Sputnik, Trung Quốc đáp lại sự kiện này bằng kế hoạch mang tầm vóc quốc gia nhằm thống trị AI vào năm 2030 và cam kết bằng tài chính để tuyên bố đó được hiện thực hoá.

Trải nghiệm tiếp xúc với AI của chúng tôi là vào năm 2012 khi một lượng nhỏ và rồi số lượng lớn các công ty AI áp dụng kỹ thuật machine- learning (học máy) tân tiến vượt bậc cho CDL. Những ứng dụng này đã mở rộng các ngành công nghiệp – phát triển các loại thuốc mới, dịch vụ khách hàng, dây chuyền sản xuất, đảm bảo chất lượng, hệ thống bán lẻ, thiết bị y tế. Công nghệ này vừa có sức ảnh hưởng mạnh mẽ vừa có tính mục đích tổng quan, nó có thể đem lại giá trị quan trọng thông quan một loạt các loại hình ứng dụng. Chúng tôi bắt đầu với mục tiêu hiểu rõ điều đó có ý nghĩa gì về mặt kinh tế. Chúng tôi biết rằng AI sẽ bị ảnh hưởng giống như nền kinh tế hay các loại công nghệ khác.

Bản chất của nó, đơn giản mà nói, rất tuyệt vời. Khi AI mới xuất hiện, nhà đầu tư mạo hiểm nổi tiếng Steve Jurvetson nói rằng: "Bất kỳ sản phẩm nào khiến bạn cảm thấy tuyệt vời như một phép màu đều sẽ được xây dựng bởi những thuật toán này."[3] Sự nhân cách hoá AI như một "phép màu" của Jurvetson đồng điệu với những câu chuyện nổi tiếng về AI trong những bộ phim như *2001: A Space Odyssey* (Du hành không gian), *Star Wars* (Chiến trang giữa các vì sao), *Blade Runner* (Tội phạm nhân bản), và những bộ phim gần đây như *Her* (Hạnh phúc ảo), *Transcendence* (Trí tuệ siêu việt), và *Ex Machina* (Cỗ máy). Chúng tôi hiểu và đồng cảm với sự nhân cách hoá những ứng dụng của AI như là "phép màu" của Jurvetson. Với tư cách là những nhà kinh tế học, công việc của chúng tôi là biến những ý tưởng có vẻ kì diệu thành đơn giản, rõ ràng và thực tiễn.

Hiểu rõ về những sự thổi phồng

Các nhà kinh tế học nhìn thế giới khác với đa số mọi người. Chúng tôi nhìn mọi thứ qua một khuôn mẫu được quyết định bởi những yếu tố như lượng cung và cầu, sản xuất và tiêu thụ, giá cả và chi phí. Tuy các nhà kinh tế thường không đồng ý với quan điểm của nhau, chúng tôi làm vậy dựa trên một khuôn mẫu chung. Chúng tôi tranh luận về những giả định và sự diễn giải, chứ không phải về những khái niệm cơ bản như vai trò của sự khan hiếm và cạnh tranh khi định giá. Cách tiếp cận này đã đem lại cho chúng tôi một điểm nhìn thuận lợi đầy độc đáo khi nhìn thế giới. Về mặt nhược điểm, quan điểm của chúng tôi khá khô khan. Còn về mặt ưu điểm, quan điểm của chúng tôi cung cấp một sự rõ ràng hữu ích cho việc đưa ra quyết định của doanh nghiệp.

Hãy bắt đầu với điều cơ bản - giá cả. Khi giá cả của một sản phẩm nào đó giảm, chúng ta mua sắm nhiều hơn. Đó là những nguyên lý kinh tế đơn giản và đang diễn ra với AI. AI ở khắp mọi nơi – trong những ứng dụng trên điện thoại của bạn, tối ưu hoá mạng lưới điện của bạn và thay thế người quản lý danh mục đầu tư chứng khoán của bạn. Có lẽ sớm thôi AI sẽ lái xe đưa đón bạn và mang các kiện hàng tới nhà bạn.

Nếu các nhà kinh tế học giỏi ở điều gì, thì đó chính là việc bỏ qua những sự thổi phồng. Khi những người khác nhìn thấy sự đổi mới, chúng tôi nhìn thấy sự giảm giá thành. Nhưng còn hơn thế nữa. Để hiểu cách AI ảnh hưởng lên tổ chức của bạn, bạn cần biết chính xác giá thành đã thay đổi ra sao và cách giá thành thay đổi sẽ ảnh hưởng tới nền kinh tế rộng lớn như thế nào. Chỉ khi đó bạn mới có thể xây dựng kế hoạch hành động. Lịch sử của nền kinh tế đã dạy cho chúng tôi rằng, ảnh hưởng của những sự đổi mới lớn thường được cảm nhận ở những nơi bất ngờ nhất.

Hãy cùng nhau xem xét câu chuyện của Internet thương mại vào năm 1995. Trong khi đa số chúng ta đều đang xem *Seinfeld*, Microsolf ra mắt Windows 95, hệ thống điều hành đa nhiệm đầu tiên của công ty. Vào cùng năm, chính phủ Hoa Kỳ đã bỏ những lệnh cấm cuối cùng để thực hiện lưu lượng truy cập thương mại trên mạng Internet, và Netscape – công ty phát minh ra trình duyệt – ăn mừng lần đầu phát hành cổ phiếu ra công chúng (IPO) của mạng Internet thương mại. Sự kiện này đánh dấu sự thay đổi khi mạng Internet chuyển từ sự tò mò về công nghệ đơn thuần sang làn sóng thương mại có thể cuốn trôi nền kinh tế.

IPO của Netscape đã khiến công ty có trị giá hơn 3 tỷ đô la, cho dù công ty chưa thực sự tạo ra lợi nhuận đáng kể nào. Các nhà đầu tư vốn mạo hiểm định giá các công ty khởi nghiệp bằng những con số triệu đô cho dù họ đang ở mức “tiền doanh thu”, đây từng là một định nghĩa mới. Những cử nhân mới tốt nghiệp với tấm bằng MBA từ chối những công việc truyền thống đầy hấp dẫn để tìm kiếm những cơ hội trên các trang web. Khi sự ảnh hưởng của mạng Internet bắt đầu lan rộng khắp các ngành nghề và làm biến đổi chuỗi giá trị, những người ủng hộ công nghệ dừng việc gọi mạng Internet là một loại hình công nghệ mới và bắt đầu đề cập tới mạng Internet như là một “Nền Kinh Tế Mới”. Mạng Internet vượt qua cả công nghệ và lan toả vào những hoạt động của con người ở mức độ quan trọng. Những chính trị gia, lãnh đạo công ty, nhà đầu tư, doanh nhân khởi nghiệp và những hãng tin tức lớn bắt đầu sử dụng thuật ngữ này. Tất cả mọi người đều bắt đầu sử dụng thuật ngữ Nền Kinh Tế Mới.

Tất cả mọi người, ngoại trừ những chuyên gia kinh tế. Chúng tôi không nhìn thấy một nền kinh tế mới. Với những chuyên gia kinh tế, nó chỉ là nền kinh tế bình thường mà thôi. Có một điều chắc chắn là một vài sự thay đổi quan trọng đã xảy ra. Hàng hoá và dịch vụ có thể được phân phối thông qua các kênh kĩ thuật số. Giao tiếp trở nên dễ dàng hơn và bạn có thể tìm thấy thông tin chỉ với cú bấm chuột vào nút tìm kiếm. Nhưng trước đây bạn cũng có thể làm được tất cả những điều này. Điều thay đổi đó là giờ bạn có thể làm điều đó với giá thành rẻ hơn. Sự phát triển của Internet mang theo sự giảm giá thành trong việc phân phối, giao tiếp và tìm kiếm. Định hình một sự phát triển công nghệ như là một sự thay đổi từ giá thành đắt đỏ tới giá thành rẻ, hoặc từ hiếm có tới dồi dào là rất quan trọng cho việc suy nghĩ về cách mà nó sẽ ảnh hưởng đến doanh nghiệp của bạn. Ví dụ, nếu bạn nhớ lại lần đầu tiên bạn sử dụng Google, bạn có thể sẽ nhớ việc đã bị mê hoặc bởi khả năng tưởng như kì diệu của việc truy cập thông tin. Theo quan điểm của một chuyên gia kinh tế, Google đã khiến việc tìm kiếm thông tin trở nên rẻ hơn. Khi giá thành của việc tìm kiếm thông tin trở nên rẻ hơn, các công ty từng kiếm được tiền bán hàng thông qua nhiều phương tiện khác nhau (ví dụ: những Trang vàng, đại lý du lịch, mẩu tin quảng cáo), cảm thấy đang phải đối mặt với những sự cạnh tranh lớn. Trong khi đó, các công ty kinh doanh phụ thuộc vào việc được người khác tìm kiếm (ví dụ, những tác giả tự xuất bản, những người bán đồ sưu tầm, những người sản xuất phim tại gia) lại trở nên thành công.

Sự thay đổi trong giá thành của một số hoạt động nhất định đã ảnh hưởng mạnh mẽ đến mô hình kinh doanh của một vài công ty và thậm chí thay đổi một số ngành công nghiệp. Tuy nhiên, luật kinh tế vẫn không thay đổi. Chúng tôi vẫn có thể hiểu tất cả mọi thứ khi nói về vấn đề cung và cầu, đồng thời vẫn có thể đặt ra chiến lược, thông báo chính sách, đồng thời hy vọng vào tương lai sử dụng những quy luật kinh tế với mô hình có sẵn.

Giá thành rẻ có ở khắp mọi nơi

Khi giá thành của một thứ gì đó mang tính thiết yếu giảm mạnh, toàn bộ thế giới có thể thay đổi. Ví dụ như ánh sáng. Có thể là bạn đang đọc cuốn sách này dưới ánh sáng nhân tạo. Hơn nữa, có thể bạn chưa bao giờ nghĩ tới việc sử dụng ánh sáng nhân tạo để đọc lại đáng quý như vậy. Ánh sáng có giá thành rẻ nên bạn có xu hướng không coi trọng nó. Nhưng, như chuyên gia kinh tế học William Nordhaus đã tỉ mỉ nghiên cứu, vào những năm đầu thế kỉ 18, bạn sẽ phải tốn gấp 400 lần số tiền bạn đang trả bây giờ để sử dụng nguồn điện tương đương.[4] Với số tiền đó, bạn sẽ chú ý hơn tới giá thành của điện và sẽ phải suy nghĩ trước khi sử dụng ánh sáng nhân tạo để đọc cuốn sách này. Sự giảm giá thành của điện đã thắp sáng thế giới. Không những điện biến ban đêm thành như ban ngày, mà nó còn giúp chúng ta sinh sống và làm việc trong những tòa nhà lớn mà ánh sáng tự nhiên khó có thể lọt vào. Chúng ta gần như không phát triển như ngày hôm nay nếu giá thành của ánh sáng nhân tạo không giảm nhiều đến vậy.

Những thay đổi trong công nghệ giúp những thứ đã từng đắt nay trở nên rẻ. Giá thành của điện giảm tới mức giờ chúng ta không còn bận tâm về việc bật nút công tắc điện. Việc giá thành giảm đáng kể như vậy tạo cơ hội cho chúng ta làm những việc trước đây không thể; nó có thể khiến những việc không thể trở thành có thể. Vậy nên, không ngạc nhiên khi những chuyên gia kinh tế học đều ám ảnh về những hệ luỵ của việc giảm giá thành đáng kể của những sản phẩm đầu vào thiết yếu như điện.

Tim Bresnahan, chuyên gia kinh tế học của Đại học Stanford và cũng là cố vấn của chúng tôi, chỉ ra rằng máy tính chỉ thực hiện các thuật toán và chỉ có vậy. Sự ra đời và thương mại hoá của máy tính khiến các thuật toán trở nên rẻ hơn.[5] Khi giá thành của các thuật toán trở nên rẻ hơn, chúng ta không những sử dụng nhiều hơn những ứng dụng thuật toán truyền thống mà chúng ta còn

sử dụng các thuật toán mới, có giá thành rẻ hơn trong những ứng dụng mà vốn không liên quan đến thuật toán, ví dụ như là âm nhạc.

Ada Lovelace là người đầu tiên nhìn thấy tiềm năng của việc trở thành lập trình viên máy tính đầu tiên. Làm việc dưới ánh điện đắt đỏ của những năm đầu thế kỷ 19, bà đã viết những chương trình được ghi nhận đầu tiên để tính toán những chuỗi số (được gọi là chuỗi số Bernoulli) cho chiếc máy tính "giả tưởng" do Charles Babbage thiết kế. Babbage cũng là một chuyên gia kinh tế học, và như bạn sẽ thấy trong cuốn sách này, đây không phải là lần duy nhất mà kinh tế học và khoa học máy tính có sự liên kết với nhau. Lovelace nhận thức được rằng thuật toán có thể - nếu nói theo ngôn ngữ khởi nghiệp hiện đại - tính toán quy mô và có khả năng làm được nhiều điều hơn thế. Bà nhận ra rằng những ứng dụng của máy tính không chỉ giới hạn trong những việc thực hành toán học: "Giả dụ, nếu những mối liên hệ cơ bản của âm thanh cao độ trong hoà âm và sáng tác âm nhạc đều dễ bị ảnh hưởng bởi những biểu hiện và những ứng dụng như vậy, thì máy móc có thể sáng tác những bản nhạc chau chuốt và khoa học ở bất kỳ mức độ khó dễ nào."[6] Khi đó máy tính chưa được phát minh, nhưng Lovelace đã hình dung ra chiếc máy sử dụng thuật toán để có thể lưu trữ và phát lại nhạc – một hình thức máy móc có tính thách thức nghệ thuật và nhân loại.

Đó chính xác là những gì đã xảy ra. Khi mà một thế kỷ rưỡi sau, giá thành của thuật toán sụt giảm và hàng ngàn ứng dụng của thuật toán mà hầu hết chưa ai từng mơ tới đã ra đời. Thuật toán là dữ liệu đầu vào quan trọng dẫn đến nhiều điều khác, khi giá thành của nó trở nên rẻ, cũng giống như điện trước đó, nó đã thay đổi thế giới. Việc nhìn mọi thứ thông qua góc nhìn giảm giá thành đã giúp chúng tôi lược bỏ những sự cường điệu hóa, tuy nhiên nó không khiến loại hình công nghệ mới nhất và tuyệt vời nhất trở nên thú vị. Bạn sẽ không bao giờ thấy cảnh Steve Jobs tuyên bố về "một loại máy móc mới", cho dù đó là những gì ông đã làm. Bằng việc giảm giá thành của một điều gì đó quan trọng, những loại máy móc mới của Jobs thực sự mang tính đột phá.

Điều đó dẫn chúng ta đến với AI. AI có tầm quan trọng về mặt kinh tế vì nó giúp một điều gì đó quan trọng trở nên rẻ hơn. Bạn có thể tưởng tượng những chú robot hoặc những sinh vật không có thật, ví dụ như những máy tính thân thiện trong *Star Trek* (Du hành giữa các vì sao), giúp bạn tránh việc nghĩ ngợi

không cần thiết. Lovelace cũng có suy nghĩ tương tự, nhưng lại nhanh chóng gạt bỏ nó. Ít nhất là khi nói về máy tính, bà đã viết, nó "không có sự căng thẳng khi làm bất kỳ việc gì. Nó có thể làm bất kỳ điều gì mà chúng ta yêu cầu nó thực hiện. Nó có thể theo dõi phân tích, nhưng nó không có khả năng dự đoán bất kỳ sự liên quan đến việc phân tích hay sự thật."[7]

Với tất cả những sự cường điệu hoá liên quan tới khái niệm AI, điều mà Alan Turing[*] sau này gọi là "Sự phản đối của quý bà Lovelace" vẫn đúng. Máy tính vẫn không có khả năng suy nghĩ, vậy nên suy nghĩ sẽ không trở nên rẻ hơn. Tuy nhiên, thứ sẽ có giá thành rẻ sẽ là một thứ có độ phổ biến rộng, giống như thuật toán, bạn sẽ không thể ngờ rằng nó lại thông dụng đến vậy và sự giảm giá thành của nó có thể ảnh hưởng tới cuộc sống của chúng ta và nền kinh tế.

[*] *Alan Mathison Turing (23/6/1912 – 7/6/1954) là một nhà toán học, logic học và mật mã học người Anh. Ông thường được xem là cha đẻ của ngành khoa học máy tính.*

Liệu còn điều gì mà AI có thể khiến giá thành trở nên rẻ hơn? *Sự dự đoán.* Vậy nên, giống như những gì mà nền kinh tế cho thấy, chúng ta không những sẽ sử dụng sự dự đoán nhiều hơn, mà chúng ta thậm chí còn ngạc nhiên khi thấy sự xuất hiện của nó ở nhiều vị trí mới.

Giá thành rẻ mang lại giá trị

Sự dự đoán là quá trình điền thông tin còn thiếu. Sự dự đoán sử dụng thông tin mà bạn có, thường được gọi là "dữ liệu", và sử dụng nó để tạo ra thông tin bạn chưa có. Nhiều cuộc thảo luận về AI nhấn mạnh sự đa dạng trong kỹ thuật dự đoán với những cái tên và nhãn hiệu không rõ ràng như: sự phân loại, sự tụ chùm, sự hồi quy, cây quyết định, ước lượng Bayes, mạng lưới nơ-ron, phân tích dữ liệu tô-pô, học sâu, học tăng cường, học sâu tăng cường, mạng lưới con nhộng… Những kỹ thuật này rất quan trọng với những chuyên gia công nghệ quan tâm tới việc triển khai AI vào một vấn đề dự đoán cụ thể. Trong cuốn sách này, chúng tôi sẽ lược bỏ những thuật toán đằng sau những phương pháp này. Chúng tôi nhấn mạnh rằng từng phương pháp đều liên quan đến dự đoán: sử dụng thông tin bạn có để tạo ra thông tin bạn chưa có. Chúng tôi tập trung vào việc giúp bạn nhận diện những tình huống mà sự dự

đoán sẽ có giá trị và cách khai thác được càng nhiều giá trị càng tốt từ dự đoán đó.

Giá thành của việc dự đoán rẻ hơn đồng nghĩa với việc sẽ có nhiều sự dự đoán hơn. Đó là nguyên lý kinh tế cơ bản: khi giá thành của thứ gì đó giảm, chúng ta có thể sử dụng nó nhiều hơn. Ví dụ, khi nền công nghiệp máy tính bắt đầu phổ biến vào những năm 1960 và giá thành của thuật toán bắt đầu giảm mạnh, chúng ta đã sử dụng nhiều thuật toán hơn trong các ứng dụng mà nó đã là thông tin đầu vào, ví dụ như tại Tổng cục điều tra Hoa Kỳ, Bộ Quốc phòng Hoa Kỳ, và NASA (gần đây đã được mô tả trong bộ phim *Hidden Figures* [tạm dịch: Số liệu ẩn]).

Tương tự vậy, sự dự đoán thường được dùng trong những công việc truyền thống, ví dụ như quản lý hàng hoá và dự báo nhu cầu. Quan trọng hơn, bởi vì giá thành của nó trở nên rẻ, nó được dùng cho những vấn đề mà trước đây chưa từng áp dụng sự dự đoán. Kathryn Howe của Integrate gọi khả năng có thể nhìn thấy vấn đề và biến nó thành vấn đề liên quan đến sự dự đoán là "sự thấu hiểu sâu sắc của AI", và ngày nay, các kỹ sư trên khắp thế giới đều đang cố gắng đạt được điều đó.

Ví dụ, chúng ta biến vấn đề giao thông thành một vấn đề liên quan đến sự dự đoán. Những chiếc xe tự động hoá đã tồn tại trong môi trường có kiểm soát hơn hai thập kỷ nay. Tuy nhiên, chúng chỉ được sử dụng ở những nơi có sơ đồ hoạt động cụ thể như nhà máy và kho hàng. Những sơ đồ tầng lầu có nghĩa là các kỹ sư có thể thiết kế robot để vận dụng nguyên lý logic "nếu-thì" đơn giản: nếu một người bước tới trước xe thì nó sẽ tự dừng lại. Nếu phần ngăn còn trống thì di chuyển qua chỗ khác. Tuy nhiên, không ai có thể sử dụng những phương tiện này trên đường phố bình thường. Có quá nhiều lo ngại có thể xảy đến – có rất nhiều "nếu" để có thể mã hoá.

Những chiếc xe tự động hoá không thể hoạt động ở ngoài môi trường có khả năng dự đoán cao và mang tính kiểm soát như vậy – cho tới khi các kỹ sư bắt đầu coi sự điều hướng như một vấn đề liên quan đến sự dự đoán. Thay vì nói với máy móc phải làm những gì trong từng trường hợp, các kỹ sư nhận ra rằng họ có thể tập trung vào từng vấn đề liên quan đến sự dự đoán: "Con người sẽ làm gì trong trường hợp đó?" Hiện giờ, các công ty đang đầu tư hàng tỷ đô la vào việc đào tạo máy móc lái xe tự động trong môi trường không kiểm soát, ngay cả trên đường phố và đường cao tốc.

Hãy thử tưởng tượng một chiếc máy AI ngồi trong xe cùng với tài xế là con người. Con người có thể lái xe hàng triệu dặm, nhận dữ liệu về môi trường thông qua mắt và tai của họ, xử lý dữ liệu bằng não bộ và rồi phản ứng lại với những dữ liệu: lái thẳng hay quẹo xe, phanh xe hay tăng tốc. Các kỹ sư sẽ cho AI đôi mắt và tai bằng cách trang bị cho xe những máy cảm biến (ví dụ như máy ảnh, ra-đa, la-de). Như vậy, AI có thể quan sát dữ liệu khi con người lái xe và đồng thời quan sát hành động của con người. Khi một dữ liệu cụ thể về yếu tố môi trường xuất hiện, con người sẽ rẽ phải, phanh, hay tăng tốc? AI càng quan sát con người nhiều, nó sẽ càng dự đoán hành động cụ thể mà người lái xe sẽ làm tốt hơn khi bị một yếu tố môi trường tác động. AI sẽ học cách lái xe thông qua việc dự đoán người lái xe sẽ làm gì trong những điều kiện cụ thể trên đường.

Quan trọng hơn, khi một dữ liệu đầu vào như sự dự đoán có giá thành rẻ, giá trị của những thứ khác cũng sẽ được nâng cao. Các chuyên gia kinh tế gọi đó là “sự bổ sung”. Giống như sự giảm giá thành của cafe sẽ khiến giá trị của đường và kem tăng, đối với những chiếc xe tự động hoá, sự giảm giá thành của sự dự đoán sẽ làm tăng giá trị của các máy cảm biến thu thập dữ liệu về môi trường xung quanh xe. Ví dụ, vào năm 2017, Intel trả hơn 15 tỷ đô la cho công ty khởi nghiệp Do Thái Mobileye, chỉ để mua công nghệ thu thập dữ liệu cho phép các phương tiện có thể nhìn thấy các vật thể (biển dừng xe, con người…) và các dấu hiệu (làn xe, đường) một cách hiệu quả.

Khi sự dự đoán có giá thành rẻ, sẽ càng có nhiều sự dự đoán và sự bổ sung liên quan đến dự đoán. Hai lực kinh tế đơn giản này dẫn đến những cơ hội mới mà máy dự đoán có thể tạo ra. Ở cấp độ thấp, máy dự đoán có thể làm giảm bớt những công tác dự đoán của con người và tiết kiệm chi phí. Khi máy bắt đầu chuyển bánh, sự dự đoán có thể thay đổi và cải thiện chất lượng của quá trình đưa ra quyết định. Nhưng ở một điểm nào đó, máy dự đoán có thể chính xác và đáng tin đến mức nó có thể thay đổi cách một tổ chức hoạt động. Một số AI có thể sẽ ảnh hưởng đến tình hình kinh tế của doanh nghiệp đến mức chúng sẽ không còn được dùng đơn thuần để nâng cao hiệu suất trong việc thực hiện chiến lược mà bản thân chúng sẽ thay đổi chiến lược.

Từ giá thành rẻ đến chiến lược

Câu hỏi mà các giám đốc điều hành thường hỏi chúng tôi đó là: “AI sẽ ảnh hưởng đến chiến lược kinh doanh của chúng tôi như thế nào?” Chúng tôi sử

dụng thí nghiệm để trả lời câu hỏi đó. Hầu hết mọi người đã quen với việc mua sắm trên Amazon. Với hầu hết các nhà bán lẻ trực tuyến, bạn sẽ truy cập trang web của họ, mua sản phẩm, rồi đặt vào giỏ hàng, trả tiền và Amazon sẽ vận chuyển chúng đến cho bạn. Hiện giờ, mô hình kinh doanh của Amazon là mua hàng-rồi-vận chuyển. Trong quá trình bạn mua hàng, AI của Amazon đề xuất những món hàng mà nó dự đoán bạn sẽ muốn mua. AI này làm việc khá hiệu quả. Tuy nhiên, nó chưa thực sự hoàn hảo. Trong trường hợp của chúng ta, AI chỉ dự đoán chính xác những gì chúng ta muốn mua vào khoảng 5%. Chúng ta thường mua khoảng một trong 20 món hàng mà nó đề xuất. Xét trên tổng số hàng triệu mặt hàng, như vậy không tệ chút nào!

Hãy tưởng tượng AI của Amazon thu thập nhiều thông tin hơn về chúng ta và sử dụng dữ liệu đó để cải thiện dự đoán của nó, một sự cải thiện giống như việc điều chỉnh âm lượng trên loa nhưng thay vì âm lượng, nó cải thiện độ chính xác của sự dự đoán của AI.

Ở một số điểm, khi nó điều chỉnh nút, độ chính xác của sự dự đoán của AI vượt ngưỡng, thay đổi mô hình kinh doanh của Amazon. Sự dự đoán trở nên đủ chính xác để Amazon thu về nhiều lợi nhuận hơn từ việc vận chuyển những món hàng nó dự đoán bạn sẽ muốn thay vì chờ bạn tự đặt hàng. Nhờ vậy, bạn sẽ không cần tới những điểm bán lẻ khác, và thực tế rằng món hàng đó có thể sẽ thúc đẩy bạn muốn mua nhiều hơn. Amazon sẽ chiếm được vị trí lớn trong ví tiền của bạn. Rõ ràng, điều này tốt cho Amazon, nhưng nó cũng sẽ tốt cho cả bạn. Amazon vận chuyển hàng tới bạn ngay trước cả khi bạn mua hàng, nếu tất cả mọi thứ đều diễn ra thuận lợi, nó sẽ giúp bạn hoàn thành toàn bộ việc mua sắm. Nút dự đoán thay đổi tăng khiến mô hình kinh doanh của Amazon từ mua hàng-rồi-vận chuyển sang vận chuyển- rồi-mua hàng.

Tất nhiên, những người mua hàng sẽ không muốn đối phó với những rắc rối của việc trả lại những món hàng họ không muốn. Vậy nên, Amazon đầu tư vào cơ sở hạ tầng cho việc trả hàng, có lẽ là những chiếc xe tải vận chuyển có tốc độ cao nhận đồ mỗi tuần một lần, thuận tiện cho việc thu thập lại những món hàng mà khách hàng không muốn.[8]

Nếu đây là mô hình kinh doanh tốt hơn, vậy sao Amazon vẫn chưa thực hiện? Bởi vì nếu bây giờ thực hiện mô hình đó, chi phí của việc thu thập và xử lý những món hàng bị trả lại sẽ lớn hơn sự tăng trưởng doanh thu từ phần lớn

của ví tiền khách hàng. Ví dụ, hôm nay chúng ta trả lại 95% số món hàng mà Amazon vận chuyển tới. Điều đó gây khó chịu cho chúng ta và tốn kém cho Amazon. Sự dự đoán không đủ tốt để Amazon thực hiện mô hình mới. Chúng ta có thể hình dung ra cảnh khi Amazon thực hiện chiến lược mới ngay trước cả khi độ chính xác của sự dự đoán đủ tốt để khiến Amazon sinh lợi nhuận vì công ty mong muốn ở một thời điểm nào đó sẽ sinh lợi nhuận. Bằng cách ra mắt sớm hơn, AI của Amazon sẽ nhận được nhiều dữ liệu sớm hơn và cải thiện nhanh hơn. Amazon nhận ra rằng họ bắt đầu càng sớm thì các đối thủ sẽ càng khó bắt kịp. Sự dự đoán tốt hơn sẽ thu hút nhiều người mua hàng hơn, nhiều người mua hàng sẽ tạo ra nhiều dữ liệu để đào tạo AI, nhiều dữ liệu sẽ dẫn đến sự dự đoán tốt hơn, và cứ thế tạo thành một chu kỳ vòng tròn. Thực hiện quá sớm có thể tốn kém, nhưng thực hiện quá muộn lại có thể nguy hiểm. [9]

Quan điểm của chúng tôi không phải là Amazon sẽ hay nên làm điều này, mặc dù những độc giả còn nghi ngờ sẽ thấy ngạc nhiên khi biết Amazon đã nhận bằng sáng chế Hoa Kỳ cho việc “dự đoán vận chuyển” vào năm 2013.10 Thay vào đó, sự thấu hiểu sâu sắc quan trọng nhất là việc dự đoán có ảnh hưởng đáng kể lên chiến lược. Trong ví dụ này, nó làm thay đổi mô hình kinh doanh của Amazon từ mua hàng-rồi-vận chuyển sang vận chuyển-rồi-mua hàng, tạo ra động lực để tích hợp theo chiều dọc vào việc vận hành dịch vụ chuyển trả hàng (bao gồm hệ thống xe tải), tăng tốc thời gian đầu tư. Tất cả những điều này đơn giản là nhờ áp dụng sự dự đoán. Điều này có ý nghĩa thế nào với chiến lược? Đầu tiên, bạn phải đầu tư thu thập thông tin tập trung vào tốc độ và khoảng cách cần vặn nút trên máy dự đoán cho lĩnh vực của bạn và những ứng dụng. Thứ hai, bạn cần phải đầu tư vào việc phát triển luận đề về những lựa chọn chiến lược được tạo ra từ việc vặn nút.

Để bắt đầu bài học “khoa học viễn tưởng” này, hãy nhắm mắt lại, tưởng tượng bạn đang đặt tay lên nút vặn của máy dự đoán, và theo những câu nói bất hủ của nhóm nhạc Spinal Tap, vặn nó tới nút mười một.

Kế hoạch cho cuốn sách

Bạn cần xây dựng nền tảng trước khi những ảnh hưởng chiến lược của máy dự đoán tới tổ chức của bạn trở nên rõ ràng. Đó chính xác là cách chúng tôi cấu trúc cuốn sách này, xây dựng một kim tự tháp từ mặt đất.

Chúng tôi xây dựng nền móng ở phần một và giải thích cách máy tự học sẽ khiến sự dự đoán trở nên tốt hơn như thế nào. Chúng tôi tiếp đến lý giải vì sao những tiến bộ mới này khác với những số liệu thống kê bạn học ở trường hoặc những chuyên gia phân tích của bạn thực hiện. Sau đó chúng tôi xem xét yếu tố bổ sung quan trọng với việc dự đoán, dữ liệu, đặc biệt là những loại dữ liệu cần thiết để có được những dự đoán tốt. Cuối cùng, chúng tôi đi sâu vào việc khi nào máy dự đoán sẽ hoạt động tốt hơn con người và khi nào con người và máy móc có thể làm việc chung để sự dự đoán chính xác hơn.

Trong phần hai, chúng tôi mô tả vai trò của sự dự đoán như là thông tin đầu vào của quá trình đưa ra quyết định và lý giải tầm quan trọng của một thành phần khác mà cộng đồng AI từ trước đến nay lãng quên: sự đánh giá. Sự dự đoán giúp việc đưa ra quyết định dễ dàng hơn bằng việc giảm sự không chắc chắn, trong khi sự đánh giá gắn giá trị cho nó.

Những vấn đề thực tiễn là trọng tâm của phần ba. Những công cụ AI khiến máy dự đoán trở nên hữu ích và là sự bổ sung của máy dự đoán được thiết kế để thực hiện một công việc cụ thể. Chúng tôi phác thảo ra ba bước có thể giúp bạn trong việc xác định khi nào xây dựng (hoặc mua) công cụ AI thì sẽ tạo ra lợi nhuận đầu tư cao nhất. Đôi khi những công cụ đó sẽ phù hợp với luồng công việc hiện có; ở một số thời điểm khác, những công cụ đó sẽ là động lực để thiết kế lại luồng công việc. Trong quá trình, chúng tôi giới thiệu công cụ trợ giúp quan trọng để cụ thể hóa những tính năng chính của công cụ AI: Canvas AI.

Chúng tôi chuyển sang nói về chiến lược ở phần thứ tư. Như chúng tôi miêu tả trong thí nghiệm Amazon, một số AI sẽ có ảnh hưởng sâu sắc đến yếu tố kinh tế của một lĩnh vực đến mức có thể thay đổi một doanh nghiệp hay một nền công nghiệp. Đó là khi AI trở thành nền tảng của chiến lược trong một tổ chức. Những AI có ảnh hưởng tới chiến lược đã khiến những nhà quản lý sản phẩm, các kỹ sư và các nhà quản lý cấp cao phải chú ý. Đôi khi thật khó để có thể nói trước khi nào một công cụ sẽ có hiệu ứng mạnh mẽ đến vậy. Ví dụ, khi sử dụng công cụ tìm kiếm của Google, một vài người đã dự đoán rằng công cụ này sẽ thay đổi ngành công nghiệp truyền thông và trở thành nền tảng của một trong những công ty có giá trị nhất trên Trái Đất.

Bên cạnh những cơ hội trên bề nổi, AI còn gây ra những lỗi hệ thống có thể ảnh hưởng đến doanh nghiệp trừ khi bạn thực hiện những hành động ngăn

chặn. Các cuộc thảo luận quan trọng dường như chỉ tập trung vào những rủi ro mà AI có thể gây ra cho nhân loại, nhưng họ ít chú ý đến những rủi ro về AI mà các tổ chức sẽ gặp phải. Ví dụ, một vài máy dự đoán được đào tạo dựa trên những dữ liệu thu thập từ con người đã “học” được những thành kiến và khuôn mẫu nguy hiểm.

Chúng tôi kết thúc cuốn sách ở phần năm bằng việc áp dụng công cụ của những chuyên gia kinh tế như chúng tôi vào những câu hỏi có tầm ảnh hưởng rộng hơn đến xã hội, xem xét năm trong số những tranh luận phổ biến về AI:

1. Liệu vẫn còn công việc? Còn.

2. Điều này sẽ gây ra sự bất bình đẳng? Có thể.

3. Sẽ có một vài công ty lớn kiểm soát mọi thứ? Phụ thuộc vào tình hình.

4. Liệu các quốc gia có tham gia vào việc hoạch định chính sách theo từng giai đoạn và liệu quyền riêng tư cũng như an ninh của chúng ta sẽ bị tước đoạt để mang lại cho những công ty trong nước lợi ích cạnh tranh tốt hơn? Một vài quốc gia sẽ làm như vậy.

5. Tận thế sắp đến rồi? Bạn vẫn còn đủ thời gian để khai thác giá trị từ cuốn sách này.

NHỮNG ĐIỂM CHÍNH

• Nền kinh tế cung cấp sự thấu hiểu rõ ràng về việc áp dụng những dự đoán giá rẻ của các doanh nghiệp. Máy dự đoán sẽ được sử dụng cho những công việc dự đoán truyền thống (hàng tồn kho và dự đoán nhu cầu) và những vấn đề mới (như điều hướng và dịch thuật). Việc giảm chi phí dự đoán sẽ ảnh hưởng đến giá trị của những thứ khác, tăng giá trị của những sự bổ sung (dữ liệu, sự đánh giá, hành động) và giảm giá trị của những sản phẩm thay thế (sự dự đoán của con người)

• Những tổ chức có thể khai thác máy dự đoán bằng cách áp dụng những công cụ AI hỗ trợ thực hiện chiến lược hiện giờ của họ. Khi những công cụ này trở nên mạnh mẽ hơn, chúng có thể thúc đẩy việc thay đổi chiến lược. Ví dụ, nếu Amazon có thể dự đoán những người mua hàng muốn gì, họ có thể

thay đổi mô hình từ mua hàng- rồi-vận chuyển sang vận chuyển-rồi-mua hàng – mang những sản phẩm tới nhà của người dùng trước khi họ đặt hàng. Sự đặt hàng này sẽ thay đổi tổ chức.

• Do những chiến lược mới mà những tổ chức sẽ theo đuổi để tận dụng ưu thế của AI, chúng ta sẽ đối mặt với nhiều sự đánh đổi mới liên quan đến cách AI sẽ ảnh hưởng đến xã hội. Những sự lựa chọn của chúng ta sẽ phụ thuộc vào nhu cầu, sở thích và sẽ gần như khác nhau giữa những quốc gia và nền văn hoá khác nhau. Chúng tôi cấu trúc cuốn sách thành năm phần để phản ánh từng lớp tác động của AI, xây dựng từ nền tảng của sự dự đoán cho đến những sự đánh đổi của xã hội: (1) Sự dự đoán, (2) Quá trình đưa ra quyết định, (3) Những công cụ, (4) Những chiến lược và (5) Xã hội

PHẦN 1SỰ DỰ ĐOÁN



3Sự kì diệu của máy dự đoán

Đ

iểm chung giữa Harry Potter, Nàng Bạch Tuyết và Macbeth là gì? Những nhân vật này đều bị thúc đẩy bởi lời tiên tri, lời dự đoán. Từ tôn giáo đến những câu chuyện cổ tích, kiến thức về tương lai là hậu quả tất yếu. Sự dự đoán ảnh hưởng đến hành vi và những quyết định.

Những người Hy Lạp cổ đại kính trọng những nhà tiên tri của họ bởi khả năng dự đoán, đôi khi là bởi những câu đố đánh lừa người hỏi. Ví dụ, vua Croesus của Lydia khi đó xem xét một cuộc tấn công mạo hiểm vào Đế quốc Ba Tư. Nhà vua không tin tưởng vào nhà tiên tri cụ thể nào, vậy nên ông quyết định kiểm tra từng người một trước khi hỏi ý kiến về việc tấn công Ba Tư. Ông gửi sứ giả đến mỗi nhà tiên tri. Vào ngày thứ 100, mỗi sứ giả được yêu cầu hỏi những nhà tiên tri về việc mà vua Croesus đang làm khi đó. Lời tiên tri ở Delphi dự đoán chính xác nhất, nên nhà vua đã hỏi và tin tưởng lời tiên tri ở đó.[1]

Trong trường hợp của vua Croesus, lời dự đoán có thể là về hiện tại. Chúng ta dự đoán liệu một giao dịch thẻ tín dụng hiện tại là hợp pháp hay bất hợp pháp, liệu một khối u trong hình ảnh y tế là ác tính hay lành tính, liệu một người nhìn vào máy ảnh của iPhone là chủ sở hữu máy hay không. Cho dù động từ gốc Latin của nó ("praedicere" – có nghĩa là "làm cho biết trước"), hiểu biết của chúng ta về mặt văn hóa của sự dự đoán nhấn mạnh khả năng nhìn thấy những thông tin ẩn giấu khác, có thể là ở quá khứ, hiện tại, hay tương lai. Quả cầu pha lê có lẽ là biểu tượng quen thuộc nhất của sự dự đoán đầy ma thuật. Chúng ta thường liên tưởng quả cầu pha lê đến những người tiên tri về sự giàu có hay đường tình duyên của ai đó trong tương lai, như trong *The Wizard of Oz* (Phù thủy xứ Oz), điều đó dẫn dắt chúng ta tới định nghĩa của dự đoán:

DỰ ĐOÁN là quá trình điền thông tin còn thiếu. Sự dự đoán lấy những thông tin mà bạn có, thường được gọi là "dữ liệu" và sử dụng chúng để tạo ra thông tin mà bạn chưa có.

Sự kì diệu của dự đoán

Nhiều năm về trước, Avi (một trong những tác giả) nhận ra một khoản giao dịch lớn, bất thường ở casino tại Las Vegas trong thẻ tín dụng của anh. Lúc đó anh ấy không ở Las Vegas. Anh ấy tới đó một lần từ rất lâu trước đó; việc thua cá độ không có sức hút đối với anh về mặt kinh tế học. Sau một cuộc trò chuyện kĩ lưỡng, nhà cung cấp thẻ tín dụng đã thay đổi được giao dịch đó và đổi thẻ cho anh ấy.

Gần đây, một vấn đề tương tự xảy ra. Ai đó đã dùng thẻ tín dụng của Avi để mua sắm. Lần này Avi không thấy khoản đó trong bản báo cáo và không phải đối mặt với quá trình giải thích với đại diện dịch vụ khách hàng lịch sự nhưng cứng rắn. Thay vào đó, anh ấy nhận được cuộc gọi trực tiếp từ phía công ty rằng thẻ của anh ấy đã bị xâm phạm và một chiếc thẻ mới đã được gửi tới. Nhà cung cấp thẻ tín dụng đã suy luận chính xác rằng giao dịch đó là lừa đảo, dựa trên thói quen tiêu tiền của Avi và nhiều dữ liệu khác. Công ty thẻ tín dụng đã tự tin vào sự dự đoán của mình đến mức họ thậm chí không chặn thẻ của anh ấy khi họ điều tra. Thay vào đó, như một phép màu, công ty gửi thẻ thay thế đến Avi mà không cần anh ấy làm bất kỳ điều gì. Tất nhiên là nhà cung cấp thẻ tín dụng không có quả cầu pha lê. Họ có dữ liệu và mô hình dự đoán tốt: máy dự đoán. Sự dự đoán tốt hơn cho phép họ ngăn chặn sự lừa đảo, như Ajay Bhalla, giám đốc mảng rủi ro doanh nghiệp và an ninh của Mastercard nói: “giải quyết vấn đề đau đầu lớn cho khách hàng về việc thẻ bị giả mạo rút tiền.”[2]

Những ứng dụng của doanh nghiệp đều phù hợp với định nghĩa về sự dự đoán của chúng tôi như là một quá trình điền thông tin còn thiếu. Sự dự đoán có ích đối với mạng lưới thẻ tín dụng để nhận biết giao dịch nào gần đây là lừa đảo. Mạng lưới thẻ tín dụng sử dụng thông tin về những giao dịch lừa đảo trước đó (và cả những giao dịch bình thường) để dự đoán rằng liệu giao dịch gần đây có phải là lừa đảo hay không. Nếu đúng là vậy, nhà cung cấp thẻ tín dụng có thể ngăn chặn những giao dịch tương lai của chiếc thẻ đó và nếu sự dự đoán được thực hiện đủ nhanh chóng, thì ngay cả những giao dịch lừa đảo trong hiện tại cũng có thể được ngăn chặn.

Khái niệm – lấy vào một loại thông tin và biến nó thành một loại thông tin khác – là trung tâm của một trong những thành tựu gần đây của AI: dịch thuật

ngôn ngữ, mục tiêu đã từ rất lâu của toàn bộ nền văn minh nhân loại, thậm chí được lưu giữ trong câu chuyện của thiên niên kỷ về Tháp Babel.* Về mặt lịch sử, cách thức tiếp cận với sự dịch thuật ngôn ngữ tự động là thuê một nhà ngôn ngữ học - chuyên gia về những quy tắc ngôn ngữ - để giải thích quy tắc và dịch chúng sang loại ngôn ngữ có thể lập trình được.[3] Đó là cách mà bạn có thể lấy một cụm từ tiếng Tây Ban Nha và, ngoài việc đơn giản là thay thế từ bằng từ, hiểu rằng bạn cần thay đổi thứ tự danh từ và tính từ để biến nó thành một câu có nghĩa.

* *Tháp Babel là một ngọn tháp to lớn được xây dựng ở thành phố Babylon, một thành phố quốc tế điển hình bởi sự hỗn tạp giữa các ngôn ngữ.*

Tuy nhiên, những tiến bộ gần đây của AI đã cho phép chúng ta định hình lại việc dịch thuật như là một vấn đề của sự dự đoán. Chúng ta có thể nhìn thấy bản chất dường như kì diệu của việc sử dụng dự đoán để dịch thuật nếu đối chiếu với sự thay đổi đột ngột về chất lượng dịch thuật của Google. Truyện gắn*The Snows of Kilimanjaro* (tạm dịch: Tuyết trên đỉnh Kilimanjaro) của Ernest Hemingway bắt đầu câu chuyện bằng hình ảnh đẹp:

Kilimanjaro là ngọn núi được bao phủ bởi tuyết cao 19.710 feet và nó được cho là ngọn núi cao nhất ở châu Phi.

Vào một ngày tháng 11 năm 2016, khi dịch bản tiếng Nhật của tập truyện ngắn kinh điển của Hemingway sang tiếng Anh nhờ Google, Giáo sư Jun Rekimoto, một nhà khoa học máy tính ở trường Đại học Tokyo đã đọc thành:

Kilimanjaro là ngọn núi cao 19.710 feet bị tuyết bao phủ và được cho là ngọn núi cao nhất ở châu Phi.

Ngày hôm sau Google dịch đã chuyển thành:

Kilimanjaro là một ngọn núi 19.710 feet được bao phủ bởi tuyết và được cho là ngọn núi cao nhất ở châu Phi.

Sự khác biệt thật đáng kinh ngạc. Chỉ qua một đêm, một bản dịch rõ ràng được dịch tự động và cứng nhắc đã trở nên vô cùng mạch lạc, từ một người có vẻ gặp khó khăn với việc dùng từ điển thành một người có vẻ như thông thạo cả hai ngôn ngữ.

Phải thừa nhận rằng, nó không thể hay bằng bản gốc của Hemingway, nhưng sự cải thiện là rất đỗi phi thường. Babel dường như đã trở lại và sự thay đổi này không phải là tình cờ hay đáng ngạc nhiên. Google đã cải thiện bộ máy đằng sau sản phẩm dịch thuật của họ để tận dụng những lợi thế gần đây của AI. Cụ thể, dịch vụ dịch thuật của Google hiện giờ phụ thuộc vào việc học sâu để dự đoán siêu nạp.

Dịch thuật từ tiếng Anh sang tiếng Nhật về cơ bản là việc dự đoán những từ và cụm từ tiếng Nhật tương thích với tiếng Anh. Thông tin còn thiếu cần được dự đoán ở đây là những cụm từ tiếng Nhật và thứ tự của chúng. Lấy vào dữ liệu từ một ngoại ngữ và dự đoán những cụm từ chính xác theo đúng thứ tự của thứ ngoại ngữ bạn biết và rồi bạn có thể hiểu được một ngôn ngữ khác. Nếu làm chính xác thì bạn có thể sẽ không nhận ra việc dịch thuật đã được thực hiện.

Các công ty đã không lãng phí thời gian của họ khi mang công nghệ kì diệu này vào sử dụng trong thương mại. Ví dụ, hơn 500 triệu người ở Trung Quốc đã sử dụng dịch vụ được ứng dụng từ việc học sâu do iFlytek phát triển để dịch thuật, phiên âm và giao tiếp bằng ngôn ngữ bản địa.

Chủ nhà sử dụng nó để giao tiếp với người thuê nhà nói ngôn ngữ khác, bệnh nhân sử dụng nó để giao tiếp với robot để chỉ dẫn, bác sĩ sử dụng nó để kê đơn chi tiết cho người bệnh và người lái xe sử dụng nó để giao tiếp với phương tiện của họ.[4]

AI càng được sử dụng nhiều sẽ càng thu nhập được nhiều dữ liệu, nó càng học hỏi nhiều hơn thì sẽ càng trở nên tốt hơn. Với nhiều người dùng như vậy, AI đang ngày càng tiến bộ nhanh chóng.

Sự dự đoán đã trở nên tốt hơn như thế nào so với trước đây?

Sự thay đổi trong Google Dịch đã minh hoạ cho cách máy tự học (trong đó học sâu là một nhánh) có thể giảm thiểu đáng kể chi phí của sự dự đoán với chất lượng đã thay đổi. Với chi phí tương tự khi nói về khả năng máy móc, Google giờ đã có thể cung cấp các bản dịch có chất lượng tốt hơn. Chi phí của việc sản xuất sự dự đoán với cùng chất lượng đã giảm một cách đáng kể.

Những sự đổi mới trong công nghệ dự đoán đang có ảnh hưởng lên những

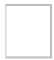
lĩnh vực mà trước nay luôn gắn liền với sự dự đoán, ví dụ như phát hiện lừa đảo. Sự phát hiện lừa đảo thẻ tín dụng đã cải thiện nhiều đến mức những công ty thẻ tín dụng phát hiện và xử lý lừa đảo trước khi chúng ta nhận ra có điều gì đó không ổn. Tuy nhiên, sự cải thiện này dường như vẫn đang dần gia tăng. Vào cuối những năm 1990, những phương pháp hàng đầu nhận diện khoảng 80% số giao dịch lừa đảo.[5] Tỉ lệ này tăng lên đến mức 90-95% vào năm 2000 và đến mức 98-99.9% hiện nay.[6] Bước nhảy cuối cùng đó là kết quả của máy tự học; sự thay đổi từ 98% đến 99.9% mang tính thay đổi lớn.

Sự thay đổi từ 98% đến 99.9% có thể mang tính gia tăng, nhưng những thay đổi nhỏ mang ý nghĩa rất lớn nếu những sai sót là vô cùng tốn kém. Sự cải thiện độ chính xác từ 85% đến 90% đồng nghĩa với việc những sai sót giảm đến 1/3. Sự cải thiện độ chính xác từ 98% đến 99.9% đồng nghĩa với việc những sai sót xuống còn 1/20.

Việc giảm giá thành dự đoán đang thay đổi nhiều hoạt động của con người. Giống như những ứng dụng đầu tiên của máy tính được áp dụng vào những vấn đề thuật toán quen thuộc như bảng điều tra dân số và thuật phóng, nhiều ứng dụng đầu tiên của dự đoán có giá thành rẻ từ máy tự học đang được áp dụng cho những vấn đề dự đoán kinh điển. Ngoài việc phát hiện lừa đảo, những ứng dụng còn được áp dụng trong các lĩnh vực như tín dụng, bảo hiểm y tế và quản lý hàng tồn kho. Tín dụng bao gồm việc dự đoán khả năng một ai đó có thể trả khoản vay. Bảo hiểm y tế bao gồm việc dự đoán một cá nhân sẽ chi trả bao nhiêu vào vấn đề chăm sóc sức khoẻ. Quản lý hàng tồn kho bao gồm việc dự đoán có bao nhiêu mặt hàng sẽ ở trong kho vào một ngày nhất định.

Gần đây, nhiều vấn đề dự đoán hoàn toàn mới xuất hiện. Nhiều trong số đó gần như bất khả thi nếu không có những tiến bộ gần đây trong công nghệ máy móc thông minh, bao gồm nhận diện đối tượng, dịch thuật ngôn ngữ và phát hiện ma tuý. Ví dụ, ImageNet Challenge là cuộc thi cao cấp thường niên dự đoán tên của một đối tượng trong một hình ảnh. Dự đoán đối tượng trong một hình ảnh có thể là một nhiệm vụ khó khăn, kể cả với con người. Dữ liệu của ImageNet lưu trữ hàng nghìn danh mục đối tượng, bao gồm nhiều giống chó và những hình ảnh tương tự. Để có thể phân biệt giống chó ngao Tây Tạng và giống chó núi Bern, hay giữa một két khoá an toàn và một tổ hợp khoá là rất khó. Con người có thể mắc sai sót khoảng 5%.[7]

Từ cuộc thi đầu tiên diễn ra vào năm 2010 cho đến cuộc thi cuối cùng vào năm 2017, khả năng dự đoán đã trở nên tốt hơn. Hình 3-1 cho thấy độ chính xác của những người chiến thắng trong cuộc thi theo từng năm. Trục dọc thể hiện tỉ lệ sai sót, vì vậy tỉ lệ càng thấp thì càng tốt. Vào năm 2010, máy dự đoán giỏi nhất mắc sai sót trong 28% số bức ảnh. Vào năm 2012, các thí sinh lần đầu tiên ứng dụng việc học sâu và tỉ lệ sai sót giảm xuống 16%. Theo đánh giá của giáo sư đồng thời là nhà khoa học máy tính Olga Russakovsky của Đại học Princeton, "2012 thực sự là năm có bước đột phá lớn trong độ chính xác, nhưng đó cũng là bằng chứng cho khái niệm của mô hình học sâu, ứng dụng đã tồn tại trong nhiều thập kỷ."[8] Sự cải thiện trong thuật toán tiếp tục diễn ra một cách nhanh chóng, và một đội đã vượt qua định mức của con người trong cuộc thi lần đầu tiên vào năm 2015. Đến năm 2017, đa số các đội trong số 38 đội đã làm tốt hơn so với định mức của con người và đội giỏi nhất đã giảm sự sai sót xuống còn một nửa. Máy móc có thể xác định những loại hình ảnh tốt hơn con người.[9]



Thế hệ AI hiện tại là cả một quá trình dài từ những máy móc thông minh của khoa học viễn tưởng. Sự dự đoán không mang lại cho chúng ta HAL từ 2001:*A Space Odyssey*, Skynet từ *The Terminator* (Kẻ hủy diệt), hay C3PO từ *Star Wars*. Nếu AI hiện nay chỉ đơn thuần là sự dự đoán, vậy thì sao nhiều người lại bàn luận về nó đến vậy? Lý do là vì sự dự đoán là thông tin đầu vào cơ bản. Bạn có thể không nhận ra nhưng sự dự đoán có mặt ở mọi nơi. Công việc kinh doanh và đời sống cá nhân của chúng ta đều đầy ắp những sự dự đoán. Thông thường sự dự đoán được ẩn dưới dạng thông tin đầu vào của quá trình đưa ra quyết định. Sự dự đoán tốt hơn đồng nghĩa với thông tin tốt hơn, cũng đồng nghĩa với quá trình đưa ra quyết định tốt hơn.

Sự dự đoán là "thông tin" để "có được thông tin hữu ích hơn"[10]. Máy dự đoán có thể tạo ra thông tin hữu ích một cách nhân tạo. Thông tin đóng vai trò vô cùng quan trọng. Sự dự đoán tốt hơn dẫn đến kết quả tốt hơn, như chúng tôi đã minh hoạ trong ví dụ phát hiện lừa đảo. Cùng với việc giá thành của sự dự đoán tiếp tục giảm, chúng tôi khám phá tính hữu ích của nó trong chuỗi đa dạng đáng kể của nhiều hoạt động khác, đồng thời cho phép thực hiện tất cả những việc, như dịch thuật ngôn ngữ bằng máy móc, mà trước đây không thể tưởng tượng được.

NHỮNG ĐIỂM CHÍNH

• Sự dự đoán là quá trình điền thông tin còn thiếu. Sự dự đoán cần thông tin mà bạn có, thường được gọi là “dữ liệu”, và sử dụng chúng để tạo ra thông tin mà bạn chưa có. Bên cạnh việc tạo ra thông tin về tương lai, sự dự đoán có thể tạo ra thông tin liên quan đến hiện tại và quá khứ. Điều này xảy ra khi sự dự đoán phân loại được những giao dịch thẻ tín dụng nào là lừa đảo, khối u nào trong hình ảnh là ác tính, hay người cầm chiếc iPhone có phải chủ sở hữu máy không.

• Ảnh hưởng của những cải thiện nhỏ trong độ chính xác của sự dự đoán có thể đánh lừa chúng ta. Ví dụ, sự cải thiện từ 85% đến 90% độ chính xác dường như gấp đôi sự cải thiện từ 98% đến 99.9% (sự gia tăng của 5% so với 2%). Tuy nhiên, sự cải thiện từ 85% đến 90% đồng nghĩa với việc sai sót giảm 1/3, trong khi đó sự cải thiện từ 98% đến 90% đồng nghĩa với việc sai sót giảm xuống còn1/20. Ở một vài trường hợp, tỉ lệ sai sót 1/20 đã mang tính chất thay đổi lớn.

• Quá trình tưởng như tầm thường của việc điền thông tin thiếu có thể khiến máy dự đoán trở nên kì diệu. Điều này đã xảy ra khi máy móc có thể nhìn (nhận dạng đối tượng), điều hướng (xe không người lái) và dịch thuật.

4Vì sao lại gọi là trí tuệ?

V

ào năm 1956, một nhóm các học giả gặp nhau ở Đại học Dartmouth, New Hampshire để vạch ra con đường nghiên cứu trí tuệ nhân tạo. Họ muốn biết liệu máy tính có thể được lập trình để có suy nghĩ nhận thức trong những việc như chơi máy tính, chứng minh định lý toán học và tương tự. Họ cũng đã suy nghĩ cẩn thận về ngôn ngữ và kiến thức để máy tính có thể mô tả đồ vật. Nỗ lực của họ bao gồm những lần thử nghiệm để cho máy tính lựa chọn và để chúng lựa chọn điều tốt nhất. Những nhà nghiên cứu rất lạc quan về khả năng của AI. Khi xin tiền tài trợ từ Quỹ Rockefeller, họ viết:

Chúng tôi sẽ nỗ lực để tìm hiểu cách khiến máy móc sử dụng ngôn ngữ, tạo ra các hình ảnh trừu tượng và khái niệm, giải quyết những vấn đề của con người, và có khả năng tự cải thiện bản thân. Chúng tôi nghĩ rằng một sự tiến bộ kinh ngạc có thể được tạo ra từ một hoặc nhiều hơn từ những vấn đề này nếu một nhóm các nhà khoa học được tuyển chọn kĩ lưỡng cùng nhau nghiên cứu về vấn đề này trong mùa hè.[1]

Chương trình nghị sự này cuối cùng đã trở thành viển vông hơn là thực tế. Trong số những thử thách khác, máy tính của những năm 1950 không đủ nhanh để có thể làm những gì các học giả hình dung.

Sau tuyên bố nghiên cứu ban đầu đó, AI cho thấy những bước tiến đầu trong việc dịch thuật, nhưng vẫn còn khá chậm. Làm việc với AI trong những môi trường cụ thể (ví dụ, một AI hoạt động như một nhà trị liệu nhân tạo) thất bại trong việc khái quát. Đầu những năm 1980 đã mang lại hy vọng rằng các kỹ sư có thể cẩn thận lập trình những hệ thống chuyên gia để nhân rộng những lĩnh vực cần tay nghề cao như chẩn đoán y khoa, nhưng những hệ thống đó rất tốn kém để phát triển, cồng kềnh và không thể giải quyết vô số những ngoại lệ và khả năng, dẫn đến khái niệm mà thường được biết đến là “mùa đông AI”.

Tuy nhiên, mùa đông cũng nhanh kết thúc.

Càng nhiều dữ liệu, mô hình càng tốt và máy tính với nhiều tính năng đã

khiến những sự phát triển gần đây trong máy tự học nâng cao sự dự đoán khả thi. Sự cải thiện trong thu nhập và lưu trữ dữ liệu khổng lồ đã cung cấp nguyên liệu cho những thuật toán học máy mới. So với những mô hình thống kê cũ, và được tạo điều kiện bởi sự phát minh của những bộ vi xử lý phù hợp hơn, những mô hình máy tự học mới đã trở nên linh hoạt đáng kể và tạo ra sự dự đoán tốt hơn – tốt đến mức một vài người đã một lần nữa miêu tả ngành khoa học máy tính này là "trí tuệ nhân tạo".

Dự đoán tỉ lệ Churn

Dữ liệu, mô hình và máy tính tốt hơn là cốt lõi của sự tiến bộ trong dự đoán. Để có thể hiểu giá trị của chúng, hãy cùng xem xét một vấn đề đã xuất hiện từ lâu của sự dự đoán: dự đoán điều mà những người làm marketing gọi là "tỉ lệ khách hàng không hài lòng". Với nhiều doanh nghiệp, thu hút khách hàng là một việc rất tốn kém và, do đó, mất khách hàng qua tỉ lệ churn là rất đắt đỏ. Một khi đã có được khách hàng, các doanh nghiệp có thể tận dụng những chi phí thu mua đó bằng việc giảm tỉ lệ churn. Trong những ngành công nghiệp dịch vụ như bảo hiểm, dịch vụ tài chính và viễn thông, quản lý tỉ lệ churn có lẽ là hoạt động marketing quan trọng nhất. Bước đầu tiên trong việc giảm tỉ lệ churn là xác định những khách hàng có khả năng đem lại rủi ro này. Các công ty có thể sử dụng công nghệ dự đoán để làm điều này.

Trong lịch sử, phương pháp cốt lõi để dự đoán tỉ lệ churn là một kỹ thuật số liệu gọi "sự hồi quy". Nghiên cứu tập trung vào cải thiện những kỹ thuật hồi quy này. Những nhà nghiên cứu đã đề xuất và thử nghiệm hàng trăm phương pháp hồi quy khác nhau trong các tạp chí học thuật và trong thực tế. Sự hồi quy có thể làm gì? Nó tìm kiếm sự dự đoán dựa trên mức trung bình của những gì xảy ra trong quá khứ. Ví dụ, nếu bạn phải xác định xem ngày mai có mưa hay không căn cứ vào tỉ lệ mưa từng ngày của tuần trước, dự đoán tốt nhất bạn có thể đưa ra sẽ chỉ ở mức trung bình. Nếu trời mưa vào hai trong số bảy ngày vừa qua, bạn có thể dự đoán rằng khả năng trời mưa ngày mai là 2/7, hay 29%. Phần lớn những gì chúng ta biết về sự dự đoán đã giúp cho những sự tính toán mức trung bình của chúng ta trở nên tốt hơn bằng cách xây dựng những mô hình có thể nhận vào nhiều dữ liệu về bối cảnh.

Chúng ta đã làm được điều này bằng việc sử dụng một thứ gọi là "mức trung bình có điều kiện". Ví dụ, nếu bạn sống ở phía Bắc California, bạn có thể có những hiểu biết nhất định trong quá khứ rằng khả năng mưa ở đây phụ thuộc

vào mùa – khả năng thấp vào mùa hè và cao vào mùa đông. Nếu bạn quan sát thấy rằng trong mùa đông, khả năng mưa của một ngày bất kỳ là 25%, thì trong mùa hè, khả năng chỉ là 5%, bạn sẽ không dự đoán rằng khả năng mưa ngày mai là ở mức trung bình – 15%. Vì sao? Vì bạn biết rằng mai là ngày mùa đông hay mùa hè, nên bạn có điều kiện để dựa vào đó và đưa ra sự đánh giá tương ứng.

Thay đổi theo mùa chỉ là một cách mà chúng ta đánh giá mức trung bình có điều kiện (cho dù đó là cách phổ biến trong thương mại bán lẻ). Chúng ta đánh giá mức trung bình có điều kiện dựa trên thời gian trong ngày, mức ô nhiễm, độ che phủ của mây, nhiệt độ đại dương, hay bất kỳ thông tin nào có sẵn.

Hoàn toàn có thể đánh giá điều kiện dựa vào nhiều yếu tố đồng thời. Liệu mai có mưa không nếu hôm nay đã mưa, trời đang vào mùa đông, trời mưa ở 200 dặm về phía Tây, trời nắng ở 100 dặm về phía Nam, mặt đất ẩm ướt, nhiệt độ Bắc Băng Dương thấp và gió thổi từ phía Tây Nam ở mức 15 dặm một giờ? Tuy nhiên, nó có thể nhanh chóng trở nên khá khó sử dụng. Riêng việc tính toán mức trung bình cho bảy loại thông tin này đã tạo ra 128 tổ hợp khác nhau. Thêm những loại thông tin khác vào sẽ tạo ra nhiều tổ hợp hơn theo cấp số nhân.

Trước khi có máy tự học, sự hồi quy đa biến đã cung cấp một cách hiệu quả để đánh giá điều kiện dựa vào nhiều yếu tố, mà không cần phải tính toán hàng chục, hàng trăm hoặc hàng nghìn mức trung bình có điều kiện. Sự hồi quy nhận dữ liệu và cố gắng tìm kiếm kết quả đồng thời hạn chế sai sót dự đoán, tối đa hoá thứ được gọi là “sự phù hợp (của mô hình hồi quy)”.

Thật may rằng thuật ngữ này chính xác về mặt toán học hơn về về mặt ngôn ngữ. Sự hồi quy giảm thiểu tối đa sai sót dự đoán ở mức trung bình và sửa lại những sai sót lớn nhiều hơn sai sót nhỏ. Đó là một phương pháp hiệu quả, đặc biệt là với những bộ dữ liệu khá nhỏ và ý thức tốt về những gì sẽ có ích trong việc dự đoán. Với tỉ lệ churn trong truyền hình cáp, nó có thể phản ánh độ thường xuyên xem TV của con người; nếu họ không sử dụng dịch vụ truyền hình cáp, vậy họ có khả năng ngừng đăng ký.

Bên cạnh đó, các mô hình hồi quy mong muốn tạo ra những kết quả không thiên vị, vậy nên với đủ những dự đoán, những dự đoán đó có thể sẽ chính

xác ở mức trung bình. Mặc dù chúng tôi thích những dự đoán không thiên vị hơn là có thiên vị (ví dụ: tự động đánh giá cao hoặc đánh giá thấp một giá trị), nhưng những dự đoán không thiên vị vẫn chưa thực sự hoàn hảo.

Việc dự đoán chính xác ở mức trung bình có thể gây ra sai sót. Sự hồi quy có thể bị lỡ mất vài feet về phía bên trái hoặc vài feet về phía bên phải. Ngay cả khi mức trung bình trùng với đáp án đúng, sự hồi quy có thể nghĩa là không bao giờ thực sự bắn trúng mục tiêu. Không giống như sự hồi quy, sự dự đoán của máy tự học có thể sai sót ở mức trung bình, nhưng khi sự dự đoán sai, chúng thường không sai quá nhiều. Những nhà thống kê mô tả điều này giống như là sự cho phép một số thiên vị để giảm phương sai.

Sự khác biệt quan trọng giữa máy tự học và phân tích hồi quy là ở cách thức các kỹ thuật mới được phát triển. Phát minh một phương pháp máy tự học mới bao gồm việc chứng minh rằng nó hoạt động tốt hơn trong thực tế. Ngược lại, phát minh một phương pháp hồi quy mới yêu cầu phải chứng minh nó hoạt động theo lý thuyết đầu tiên. Việc tập trung vào hoạt động trong thực tế giúp những nhà sáng chế máy tự học có không gian để thử nghiệm, ngay cả khi phương pháp của họ tạo ra những ước tính chính xác ở mức trung bình hoặc mang tính thiên vị. Sự tự do trong thực nghiệm này đã thúc đẩy nhanh chóng sự phát triển bằng cách tận dụng sự phong phú dữ liệu và máy tính xử lý nhanh.

Trong suốt cuối những năm 1990 và đầu những năm 2000, những thí nghiệm với máy tự học để dự đoán tỉ lệ churn của khách hàng đã có những thành công nhất định. Những phương pháp của máy tự học đang được cải thiện, nhưng sự hồi quy nhìn chung vẫn hoạt động tốt hơn. Dữ liệu không đủ phong phú và máy tính không đủ tốt để tận dụng lợi thế mà máy tự học có thể làm. Ví dụ, Trung tâm Teradata của Đại học Duke đã tổ chức giải đấu khoa học dữ liệu vào năm 2004 để dự đoán tỉ lệ churn. Những giải đấu như vậy khi đó không phổ biến. Ai cũng có thể dự thi và những bài dự thi chiến thắng sẽ nhận được giải thưởng tiền mặt. Những bài dự thi chiến thắng sử dụng mô hình hồi quy. Một vài phương pháp máy tự học đã được thể hiện một cách tương đối tốt, nhưng những phương pháp mạng lưới nơ-ron mà sau này sẽ thúc đẩy cách mạng AI đã không thể hiện tốt. Cho đến năm 2016, mọi thứ đã thay đổi. Những mô hình tỉ lệ churn tốt nhất sử dụng máy tự học và những mô hình học sâu (mạng lưới nơ-ron) nhìn chung thể hiện vượt trội hơn những

mô hình khác.

Điều gì đã thay đổi? Đầu tiên, dữ liệu và máy tính cuối cùng đã đủ tốt để cho phép máy tự học chiếm ưu thế. Vào những năm 1990, thật khó để có thể xây dựng bộ dữ liệu đủ lớn. Ví dụ, một nghiên cứu kinh điển về dự đoán tỉ lệ churn sử dụng 650 khách hàng và ít hơn 30 biến. Cho đến năm 2004, bộ xử lý và lưu trữ máy tính đã được cải thiện. Trong giải đấu Duke, tập dữ liệu đào tạo lưu trữ thông tin của hàng trăm biến cho hàng chục nghìn khách hàng. Với những biến và khách hàng được bổ sung thêm này, những phương pháp máy tự học bắt đầu được thực hiện, thậm chí là tốt hơn hồi quy.

Hiện giờ những nhà nghiên cứu dự đoán tỉ lệ churn dựa vào hàng nghìn biến và hàng triệu khách hàng. Sự cải thiện trong sức mạnh máy tính đồng nghĩa với việc có thể sử dụng lượng lớn dữ liệu, bao gồm văn bản, hình ảnh và những con số. Ví dụ, trong mô hình tỉ lệ churn ở điện thoại, những nhà nghiên cứu tận dụng dữ liệu của những cuộc gọi từng giờ bên cạnh những biến tiêu chuẩn như kích thước hoá đơn và thanh toán đúng hạn.

Những phương pháp máy tự học cũng trở nên tốt hơn ở việc tận dụng những dữ liệu có sẵn. Trong cuộc thi Duke, yếu tố quan trọng của thành công là lựa chọn trong số hàng trăm biến có sẵn và lựa chọn mô hình thống kê nào để sử dụng. Các phương pháp tốt nhất ở thời điểm đó, cho dù là máy tự học hay sự hồi quy kinh điển, đều sử dụng sự kết hợp của trực giác và những bài kiểm tra thống kê để lựa chọn các biến và mô hình. Hiện giờ, những phương pháp máy tự học, và đặc biệt là những phương pháp học sâu, đều cho phép tính linh hoạt trong mô hình và điều này có nghĩa là các biến có thể kết hợp với nhau theo những cách bất ngờ. Những tổ hợp đó rất khó để dự đoán, nhưng chúng có thể hỗ trợ sự dự đoán rất nhiều. Bởi vì chúng rất khó để có thể đoán trước, những người lập mô hình sẽ không sử dụng chúng khi dự đoán với những kỹ thuật hồi quy tiêu chuẩn. Máy tự học đưa ra những lựa chọn mà sự kết hợp và tương tác có thể quan trọng đối với máy móc, chứ không phải với lập trình viên.

Sự cải thiện trong phương pháp máy tự học nói chung và học sâu nói riêng, có nghĩa rằng khả năng biến những dữ liệu có sẵn thành những dự đoán tỉ lệ churn chính xác là hoàn toàn có thể. Và những phương pháp máy tự học hiện giờ rõ ràng đã chiếm ưu thế hơn so với sự hồi quy và những kỹ thuật khác.

Vượt xa hơn tỉ lệ churn

Máy tự học đang cải thiện sự dự đoán trong vô số những bối cảnh khác ngoài tỉ lệ churn, từ thị trường tài chính đến thời tiết.

Cuộc khủng hoảng tài chính vào năm 2008 là một thất bại đáng kinh ngạc của những phương pháp dự đoán dựa trên sự hồi quy. Một yếu tố thúc đẩy khủng hoảng tài chính là những dự đoán về khả năng mặc định của những nghĩa vụ nợ thế chấp, hay được gọi là CDO. Vào năm 2007, những công ty xếp hạng như Standard & Poor's dự đoán rằng những CDO hạng AAA có ít hơn 1/800 cơ hội bị thất bại trong việc không đạt được lợi nhuận trong vòng năm năm. Năm năm sau, hơn một trong bốn CDO thất bại trong việc đạt được lợi nhuận. Dự đoán ban đầu hoàn toàn sai cho dù đã có rất nhiều dữ liệu mặc định trước đây.

Sự thất bại không phải do thiếu dữ liệu, mà thay vào đó là cách những chuyên gia phân tích sử dụng dữ liệu để dự đoán. Những công ty xếp hạng dự đoán dựa trên những mô hình hồi quy giả định rằng giá nhà trong những thị trường khác nhau thì thường không liên quan đến nhau. Điều đó hoá ra là sai, không chỉ vào năm 2007 và cả trước đó cũng vậy. Bao gồm cả khả năng một cú sốc có thể gây ảnh hưởng nhiều thị trường nhà cùng một lúc và khả năng này gia tăng khi bạn mất CDO, ngay cả khi chúng được phân phối khắp nhiều thành phố của Hoa Kỳ.

Những chuyên gia phân tích xây dựng mô hình hồi quy của họ trên những giả thuyết về những điều họ tin là quan trọng và niềm tin là không cần thiết với máy tự học. Những mô hình máy tự học đặc biệt tốt trong việc xác định những biến nào có thể hoạt động tốt nhất và nhận ra một vài điều không quan trọng trong khi đó những điều khác lại quan trọng một cách đáng ngạc nhiên. Hiện giờ, trực giác và giả thuyết của chuyên gia phân tích trở nên ít quan trọng hơn. Bằng cách này, máy tự học dự đoán dựa trên những mối liên quan có thể không được dự đoán trước, bao gồm giá nhà ở Las Vegas, Phoenix và Miami có thể thay đổi cùng một lúc.

Nếu nó chỉ là sự dự đoán, thì sao lại gọi là "trí tuệ"?

Những tiến bộ gần đây trong máy tự học đã thay đổi cách chúng ta sử dụng số liệu thống kê để dự đoán. Thật hấp dẫn khi xem những phát triển gần đây

nhất trong AI và máy tự học như là “số liệu thống kê truyền thống được tiêm steroid*”. Theo một cách nào đó thì điều đó là sự thật, bởi vì mục tiêu cuối cùng là tạo ra một sự dự đoán để điền thông tin còn thiếu. Hơn nữa, quá trình máy tự học bao gồm sự tìm kiếm giải pháp có xu hướng làm giảm thiểu sai sót.

* *Steroid là thuốc tăng cơ bắp, có ảnh hưởng rất lớn tới quá trình hóa học của cơ thể, kích thích tăng trưởng và các chức năng sinh lý khác.*

Vậy điều gì khiến máy tự học trở thành một công nghệ máy tính mang tính chuyển đổi xứng đáng với cái tên “trí tuệ nhân tạo”? Trong một số trường hợp, sự dự đoán tốt đến mức chúng ta có thể dùng sự dự đoán thay cho logic dựa trên quy tắc.

Sự dự đoán hiệu quả thay đổi cách máy tính được lập trình. Những phương pháp thống kê truyền thống hay thuật toán của mệnh đề nếu-thì đều không hoạt động tốt trong những môi trường phức tạp. Bạn muốn nhận diện một con mèo trong nhóm những bức ảnh? Cụ thể là những chú mèo đều có nhiều màu sắc và tư thế khác nhau. Chúng có thể đứng, ngồi, nằm, nhảy nhót, hoặc trông cau có. Chúng có thể ở trong hay ở ngoài. Điều này sẽ nhanh chóng trở nên phức tạp. Do đó, ngay cả khi thực hiện một công việc dễ dàng cũng đều đòi hỏi sự cẩn thận và đó chỉ là ví dụ với những chú mèo. Vậy nếu chúng ta muốn một cách miêu tả tất cả những đối tượng trong một bức ảnh thì sao? Chúng ta cần phải miêu tả riêng từng đối tượng.

Một công nghệ quan trọng, nền tảng cho những tiến bộ gần đây, được gọi là “học sâu”, dựa vào cách tiếp cận gọi là “tuyên truyền ngược”. Nó tránh được tất cả những điều phức tạp ở trên, giống như não bộ con người, bằng cách học hỏi qua những ví dụ (liệu những nơ-ron nhân tạo bắt chước những nơ-ron thật hay không là một sự phân tâm thú vị khỏi tính ứng dụng của loại công nghệ này).

Nếu bạn muốn một đứa trẻ biết từ “con mèo”, thì mỗi khi bạn thấy con mèo, hãy nói từ này. Điều đó cũng tương tự với máy tự học. Bạn cho nó một số bức ảnh của mèo với nhãn dán “mèo” và một số bức ảnh không có mèo mà không có nhãn dán “mèo”. Máy sẽ học cách nhận diện những mẫu pixel liên quan đến nhãn dán “mèo”. Nếu bạn có những hình ảnh về cả mèo và chó thì

mối liên hệ giữa mèo và đối tượng bốn chân sẽ được củng cố. Không cần phải cụ thể hoá hơn, một khi bạn cho máy vô số triệu hình ảnh với nhiều biến khác nhau (bao gồm những hình ảnh không có chó) và dán nhãn vào máy, nó sẽ phát triển nhiều hơn những mối liên hệ và học hỏi để phân biệt giữa mèo và chó. Nhiều vấn đề đã chuyển từ vấn đề thuật toán ("con mèo có những đặc điểm gì?") sang vấn đề dự đoán ("liệu hình ảnh không có nhãn dán này có đặc điểm giống với những con mèo mà tôi đã thấy trước đây?"). Máy tự học sử dụng những mô hình xác suất để giải quyết vấn đề.

Vậy tại sao nhiều chuyên gia công nghệ gọi máy tự học là "trí tuệ nhân tạo"? Bởi vì thông tin đầu ra của máy tự học – sự dự đoán – là thành phần quan trọng của trí tuệ, độ chính xác của sự dự đoán cải thiện qua sự học hỏi, và độ chính xác cao của sự dự đoán cho phép máy thực hiện những công việc mà cho đến giờ, thường gắn liền với trí tuệ con người, ví dụ như nhận diện đối tượng.

Trong cuốn sách *On Intelligence* (tạm dịch: Bàn về trí tuệ), Jeff Hawkins là một trong những người đầu tiên cho rằng sự dự đoán là cơ sở cho trí tuệ con người. Bản chất lý thuyết của ông là trí tuệ con người, cốt lõi của sự sáng tạo và gia tăng năng suất, là do cách não bộ của chúng ta sử dụng ký ức để đưa ra sự dự đoán: "Chúng ta đang liên tục đưa ra những sự dự đoán song song ở mức độ thấp trong khắp các giác quan. Nhưng đó không phải là tất cả. Tôi cho rằng nó có một vai trò mạnh mẽ hơn nhiều. Sự dự đoán không chỉ là một trong số những điều mà não bộ bạn thực hiện. Đó là chức năng chính của tân vỏ não và là nền tảng của trí tuệ. Vỏ não chính là cơ quan dự đoán."[2]

Tác phẩm của Hawkins gây ra khá nhiều tranh cãi. Những ý tưởng của ông được tranh luận trong các tài liệu về tâm lý học và nhiều nhà khoa học máy tính bác bỏ sự nhấn mạnh của ông về việc vỏ não là mô hình cho máy dự đoán. Ý kiến rằng AI có thể vượt qua bài kiểm tra Turing (máy móc có khả năng đánh lừa con người tin rằng cỗ máy đó thực sự là con người) vẫn còn khá xa với thực tế. Thuật toán AI hiện tại không thể lý giải, và hơn nữa rất khó chất vấn chúng để hiểu rõ nguồn gốc sự dự đoán của chúng.

Không có sự phân biệt liệu mô hình nền tảng đó có phù hợp hay không, sự nhấn mạnh của Hawkins rằng sự dự đoán là nền tảng của trí tuệ thực sự hữu ích cho việc hiểu rõ về ảnh hưởng của những thay đổi gần đây trong AI. Ở

đây chúng tôi nhấn mạnh đến hệ quả của những cải thiện đáng kể trong công nghệ dự đoán. Rất nhiều nguyện vọng của các học giả tại hội nghị Dartmouth năm 1956 nay đã trong tầm với. Theo nhiều cách khác nhau, máy dự đoán có thể "sử dụng ngôn ngữ, tạo ra các hình ảnh trừu tượng và khái niệm, giải quyết những vấn đề mà hiện giờ [lúc bấy giờ là năm 1955] chỉ có con người thực hiện được và tự cải thiện bản thân."[3]

Chúng tôi không suy đoán liệu quá trình này có dự đoán trước sự xuất hiện của trí tuệ nhân loại, "điểm dị biệt", hay Skynet. Tuy nhiên, như bạn sẽ thấy, sự tập trung vào sự dự đoán vẫn sẽ gợi ý những thay đổi phi thường trong vài năm tới. Giống như cách các thuật toán trở nên rẻ hơn nhờ máy tính đã chứng minh hiệu quả mạnh mẽ bằng việc đem đến những thay đổi đáng kể trong kinh doanh và đời sống cá nhân, những sự thay đổi tương tự sẽ xảy ra nhờ có sự dự đoán có giá thành rẻ.

Nhìn chung, cho dù nó là trí tuệ hay không, sự phát triển từ lập trình máy tính quyết định đến xác suất là một bước chuyển tiếp chức năng quan trọng, phù hợp với quá trình phát triển của khoa học xã hội và vật chất. Nhà triết học Ian Hacking, trong cuốn sách The Taming of Chance (tạm dịch: Làm chủ cơ hội), nói rằng, trước thế kỷ 19, xác suất chính là lãnh thổ của những người đánh cược.[4] Cho đến thế kỷ 19, dữ liệu điều tra dân số đã tăng lên nhờ sử dụng kỹ thuật toán học mới xuất hiện là xác suất vào khoa học xã hội. Thế kỷ 20 chứng kiến sự sắp xếp lại hiểu biết mang tính quan trọng của chúng ta về thế giới vật chất, chuyển từ Thuyết quyết định của Newton (tất cả các sự việc xảy ra là do những điều tất yếu và do đó không thể tránh được) sang Nguyên lý bất định của cơ học lượng tử. tiến bộ quan trọng nhất trong khoa học máy tính của thế kỷ 21 phù hợp với những tiến bộ về khoa học xã hội và vật chất sau: sự công nhận rằng thuật toán hoạt động tốt nhất khi xác suất được cấu trúc dựa trên dữ liệu.

NHỮNG ĐIỂM CHÍNH

• Khoa học về máy tự học có những mục tiêu khác so với thống kê số liệu. Trong khi thống kê số liệu nhấn mạnh vào việc chính xác ở mức trung bình, máy tự học không cần điều đó. Thay vào đó, mục tiêu của máy tự học là hoạt động hiệu quả. Sự dự đoán vẫn sẽ được ưu tiên miễn là chúng tốt hơn (điều trở nên khả thi nhờ sức mạnh của máy tính). Điều này đem lại cho những nhà

khoa học sự tự do để thực nghiệm và thúc đẩy sự cải thiện nhanh chóng bằng cách tận dụng lợi thế của dữ liệu phong phú và sự xuất hiện của máy tính nhanh trong hơn một thập kỷ qua.

• Những phương pháp thống kê số liệu truyền thống yêu cầu sự chính xác của giả thuyết hoặc ít nhất là trực giác của con người đối với mô hình cụ thể. Máy tự học có ít nhu cầu xác định cái gì sẽ đi vào mô hình và có thể chứa lượng tương đương những mô hình phức tạp hơn với những tương tác giữa các biến.

• Những tiến bộ gần đây trong máy tự học thường được gọi là những tiến bộ trong trí tuệ nhân tạo vì: (1) những hệ thống dự đoán dựa trên kỹ thuật này học hỏi và cải thiện theo thời gian; (2) những hệ thống này tạo ra những dự đoán quan trọng chính xác hơn những cách tiếp cận khác dưới những điều kiện cụ thể, và một vài chuyên gia cho rằng sự dự đoán là trung tâm của trí tuệ; và (3) độ chính xác cao trong dự đoán của những hệ thống này cho phép chúng thực hiện những công việc, ví dụ như dịch thuật và điều hướng, mà trước đây thường được cho là lĩnh vực độc quyền của trí tuệ con người. Kết luận của chúng tôi không dựa trên ý kiến cá nhân về việc liệu những tiến bộ trong sự dự đoán có đại diện cho những tiến bộ về trí tuệ hay không. Chúng tôi tập trung vào hệ quả của sự giảm giá thành của sự dự đoán, không phải sự giảm giá thành của trí tuệ.

5Dữ liệu là nguyên liệu mới

H

al Varian, chuyên gia kinh tế trưởng tại Google, nói vào năm 2013, “Vào một tỷ giờ trước, con người hiện đại bắt đầu xuất hiện. Vào một tỷ phút trước, Cơ Đốc giáo bắt đầu. Vào một tỷ giây trước, máy tính IBM được phát hành. Một tỷ kết quả tìm kiếm trên Google… xảy ra trong sáng nay.”[1]

Google không phải là công ty duy nhất có lượng dữ liệu khổng lồ. Từ những công ty lớn như Facebook và Microsoft đến chính quyền địa phương và những công ty khởi nghiệp, việc thu nhập dữ liệu đã trở nên rẻ hơn và dễ dàng hơn bao giờ hết. Dữ liệu này đều có giá trị. Hàng tỷ kết quả tìm kiếm đồng nghĩa với việc Google có thể cải thiện dịch vụ của họ với hàng tỷ dòng dữ liệu. Một số người gọi loại dữ liệu này là “nguồn nguyên liệu mới”.

Máy dự đoán dựa vào dữ liệu. Càng nhiều dữ liệu tốt dẫn đến những kết quả dự đoán tốt hơn. Về mặt kinh tế, dữ liệu là một sự bổ sung quan trọng cho sự dự đoán. Nó trở nên có giá trị khi sự dự đoán có giá thành rẻ hơn.

Với AI, dữ liệu đóng ba vai trò. Đầu tiên là *dữ liệu đầu vào,* dữ liệu này được cấp cho thuật toán và sử dụng để tạo ra sự dự đoán. Thứ hai là dữ liệu đào tạo, được dùng để tạo ra thuật toán đầu tiên. Dữ liệu đào tạo được sử dụng để đào tạo AI trở nên đủ tốt để dự đoán. Cuối cùng là dữ liệu phản hồi, được dùng để cải thiện hiệu suất của thuật toán bằng kinh nghiệm. Trong một số trường hợp, nhiều sự chồng chéo xuất hiện mà ở đó những dữ liệu đóng cả ba vai trò.

Nhưng để thu được dữ liệu có thể rất tốn kém. Bởi vậy, khoản đầu tư bao gồm sự đánh đổi giữa lợi ích của việc có nhiều dữ liệu hơn và chi phí để có được nó. Để đưa ra những quyết định đầu tư dữ liệu đúng đắn, bạn phải hiểu cách mà máy dự đoán sử dụng dữ liệu.

Sự dự đoán yêu cầu dữ liệu

Trước khi có nhiều sự quan tâm đối với AI, đã từng có nhiều người hứng thú với dữ liệu lớn. Sự đa dạng, số lượng và chất lượng của dữ liệu đã tăng đáng

kể trong vòng 20 năm qua. Hình ảnh và văn bản hiện giờ ở dạng kỹ thuật số, vậy nên máy móc có thể phân tích được chúng. Máy cảm biến có mặt ở khắp mọi nơi. Sự quan tâm được đánh giá dựa trên khả năng của dữ liệu trong việc giúp con người giảm thiểu sự không chắc chắn và hiểu thêm về những gì đang xảy ra.

Hãy cân nhắc những máy cảm biến đã được cải tiến có khả năng theo dõi nhịp tim của con người. Rất nhiều công ty và tổ chức phi lợi nhuận với những cái tên nghe có vẻ liên quan đến y tế như AliveCor và Cardiio đang xây dựng những sản phẩm sử dụng dữ liệu nhịp tim. Ví dụ, công ty khởi nghiệp Cardiogram cung cấp một ứng dụng iPhone sử dụng dữ liệu nhịp tim từ Apple Watch để tạo ra khối lượng thông tin lớn: thước đo nhịp tim theo từng giây cho tất cả những ai sử dụng ứng dụng. Người dùng có thể nhìn thấy khi nào nhịp tim của họ tăng đột biến trong ngày hoặc nhịp tim của họ tăng hay giảm trong suốt một năm, thậm chí trong suốt một thập kỷ.

Nhưng sức mạnh tiềm năng của những sản phẩm như vậy xuất phát từ sự kết hợp của dữ liệu phong phú với máy dự đoán. Cả những nhà nghiên cứu học thuật và những nhà nghiên cứu ngành công nghiệp đều đã chỉ ra rằng điện thoại thông minh có thể dự đoán những nhịp tim bất thường (nói theo thuật ngữ y khoa, sự rung tâm nhĩ).[2] Vậy, với máy dự đoán, những sản phẩm mà Cardiogram, AliveCor, Cardiio và những công ty khác đang xây dựng, đều sử dụng dữ liệu nhịp tim để giúp chẩn đoán bệnh tim. Cách tiếp cận chung là sử dụng dữ liệu nhịp tim để dự đoán những thông tin chưa biết về liệu người dùng có nhịp tim bất thường hay không.

Dữ liệu đầu vào này là cần thiết để vận hành máy dự đoán. Do máy dự đoán không thể chạy mà không có dữ liệu đầu vào, trái ngược với dữ liệu đào tạo và dữ liệu phản hồi. Người dùng chưa có kinh nghiệm không thể nhìn thấy mối liên hệ giữa dữ liệu nhịp tim và nhịp tim bất thường từ dữ liệu thô. Ngược lại, Cardiogram có thể phát hiện nhịp tim bất thường với 97% độ chính xác bằng việc sử dụng hệ thống mạng nơ-ron sâu của họ.[3]

Những sự bất thường như vậy gây ra khoảng 1/4 khả năng đột quỵ. Với sự dự đoán tốt hơn, các bác sĩ có thể điều trị tốt hơn. Một số loại thuốc nhất định có thể ngăn ngừa đột quỵ. Để có thể làm được như vậy, cá nhân những người dùng cần phải cung cấp dữ liệu nhịp tim của họ. Nếu không có dữ liệu cá

nhân, máy không thể dự đoán được rủi ro cho người đó. Sự kết hợp của máy dự đoán với dữ liệu cá nhân của người dùng giúp dự đoán tốt hơn về khả năng của một người có nhịp tim bất thường.

Cách máy học hỏi từ dữ liệu

Thế hệ công nghệ AI hiện nay được gọi là "máy tự học" cũng có lý do. Máy móc học hỏi từ dữ liệu. Trong trường hợp của máy đo nhịp tim, để có thể dự đoán nhịp tim bất thường (và khả năng của việc đột quỵ) từ dữ liệu nhịp tim, máy dự đoán phải học mối liên hệ giữa dữ liệu với tỷ lệ thực tế của nhịp tim bất thường. Để có thể làm được vậy, máy dự đoán cần kết hợp dữ liệu đầu vào từ Apple Watch – điều mà những nhà thống kê gọi là "biến độc lập" – với thông tin về nhịp tim bất thường ("biến phụ thuộc").

Để máy dự đoán có thể học được, thông tin về nhịp tim bất thường phải đến từ cùng một người với dữ liệu nhịp tim được thu thập bởi Apple Watch. Vậy nên máy dự đoán cần dữ liệu từ nhiều người có nhịp tim bất thường, cùng với dữ liệu nhịp tim sẵn có của họ. Quan trọng là nó cũng cần dữ liệu từ những người không có nhịp tim bất thường cùng với dữ liệu nhịp tim của những người đó. Máy dự đoán sau đó sẽ so sánh những mẫu nhịp tim với nhịp điệu bình thường và bất thường. Sự so sánh này sẽ dẫn đến sự dự đoán. Nếu mẫu dữ liệu nhịp tim mới giống với mẫu "đào tạo" của những người có nhịp tim bất thường hơn là với mẫu của những người có nhịp tim bình thường, thì máy sẽ dự đoán rằng bệnh nhân này có nhịp tim bất thường.

Giống như nhiều ứng dụng y khoa khác, Cardiogram thu thập dữ liệu bằng cách làm việc với nhiều nhà nghiên cứu học thuật đã theo dõi 6.000 người dùng để hỗ trợ việc nghiên cứu. Trong số 6.000 người dùng, khoảng 200 người đã được chẩn đoán với nhịp tim bất thường. Vậy tất cả những gì Cardiogram làm là thu nhập dữ liệu về các mẫu nhịp tim từ Apple Watch và so sánh.

Những sản phẩm như vậy sẽ tiếp tục cải thiện độ chính xác của sự dự đoán ngay cả sau khi được phát hành. Máy dự đoán cần dữ liệu phản hồi xem liệu sự dự đoán của nó có chính xác hay không. Vậy nên, nó cần dữ liệu về tỷ lệ của nhịp tim bất thường trong số những người dùng. Máy sẽ kết hợp dữ liệu nhịp tim bất thường với dữ liệu đầu vào về việc theo dõi tim mạch để cung cấp phản hồi và liên tục cải thiện độ chính xác của sự dự đoán.

Tuy nhiên, sự thu thập dữ liệu đào tạo có thể sẽ là một thách thức. Để dự đoán những đối tượng trong cùng một nhóm (trong trường hợp này, bệnh nhân mắc bệnh tim), bạn cần thông tin về tỉ lệ kết quả đầu ra cũng như thông tin hữu ích cho việc dự đoán kết quả đầu ra trong bối cảnh mới (theo dõi tim mạch). Điều này đặc biệt khó khăn khi dự đoán là về sự kiện trong tương lai. Để đưa ra sự dự đoán này, bạn cần dữ liệu ở thời điểm bạn cần đưa ra sự dự đoán.

Rất nhiều ứng dụng AI thương mại có cấu trúc như sau: sử dụng sự kết hợp của dữ liệu đầu vào và kết quả đầu ra ước tính để tạo ra máy dự đoán, và rồi sử dụng dữ liệu đầu vào từ một tình huống mới để dự đoán kết quả đầu ra của tình huống đó. Nếu bạn có thể thu thập dữ liệu từ kết quả đầu ra đó, máy dự đoán của bạn có thể học hỏi liên tục thông qua phản hồi.

Quyết định liên quan đến dữ liệu

Dữ liệu thường sẽ rất tốn kém để có được, nhưng máy dự đoán không thể hoạt động mà không có nó. Máy dự đoán cần dữ liệu để tạo ra, hoạt động và cải thiện.

Do đó, bạn phải đưa ra quyết định về quy mô và phạm vi của việc thu thập dữ liệu. Bạn cần bao nhiêu loại dữ liệu khác nhau? Cần bao nhiêu đối tượng khác nhau để đào tạo máy? Tần suất bạn cần thu nhập dữ liệu? Càng nhiều loại, càng nhiều đối tượng, tần suất càng lớn đồng nghĩa với giá thành càng cao nhưng càng có khả năng thu lợi nhuận lớn. Khi suy nghĩ về quyết định này, bạn cần phải cẩn thận xác định điều bạn muốn dự đoán. Vấn đề dự đoán cụ thể sẽ nói cho bạn biết bạn cần cái gì.

Cardiogram muốn dự đoán số lần đột quỵ. Họ sử dụng nhịp tim bất thường như là một sự đại diện (đã được chứng nhận về mặt y khoa).4 Một khi họ đặt ra mục tiêu dự đoán, họ chỉ cần dữ liệu nhịp tim của mỗi người dùng ứng dụng. Họ có thể cần thông tin về giấc ngủ, hoạt động thể chất, gia đình, bệnh sử và tuổi tác. Sau khi hỏi một vài câu hỏi để thu nhập thông tin về tuổi tác và những thông tin khác, họ chỉ cần một thiết bị để đo lường chuẩn nhịp tim.

Cardiogram cũng cần dữ liệu để đào tạo – 6.000 người trong hệ thống dữ liệu đào tạo của họ, một phần trong số đó có nhịp tim bất thường. Mặc dù có nhiều loại máy cảm biến và nhiều chi tiết khác nhau về người dùng có thể có

sẵn, Cardiogram chỉ cần thu nhập số lượng ít thông tin về hầu hết người dùng của họ. Và họ chỉ cần thông tin về nhịp tim bất thường từ những người mà họ dùng để đào tạo máy AI của họ. Bằng cách này, số biến là tương đối nhỏ.

Để có thể đưa ra sự dự đoán tốt, máy cần có đủ cá thể (hoặc đơn vị phân tích) trong dữ liệu đào tạo. Số cá thể cần có phụ thuộc vào hai yếu tố: đầu tiên, độ nhạy của "tín hiệu" với "tiếng ồn", và thứ hai, độ chính xác của sự dự đoán để trở nên hữu dụng. Hay nói cách khác, số lượng cá thể cần có phụ thuộc vào việc nhịp tim là yếu tố dự đoán nhịp tim bất thường mạnh hay yếu và sự tốn kém mà sai sót có thể gây ra. Nếu nhịp tim là yếu tố dự đoán mạnh và sai sót không quan trọng, vậy thì chúng ta chỉ cần một vài người. Nếu nhịp tim là yếu tố dự đoán yếu và mỗi sai sót có thể gây nguy hiểm cho tính mạng, thì chúng ta cần hàng nghìn thậm chí là hàng triệu cá thể. Cardiogram đã sử dụng thông tin của 6.000 người trong nghiên cứu sơ bộ của họ, bao gồm chỉ 200 người với nhịp tim bất thường. Theo thời gian, một cách để thu thập thêm dữ liệu là từ phản hồi của việc liệu những người sử dụng ứng dụng có nhịp tim bất thường hay không.

Vậy con số 6.000 từ đâu ra? Những nhà khoa học dữ liệu có những công cụ tuyệt vời để đánh giá khối lượng dữ liệu cần có để thu được sự dự đoán đáng tin và chính xác. Những công cụ này được gọi là "tính toán công suất" và chúng sẽ nói cho bạn biết cần bao nhiêu đơn vị phân tích để cho ra sự dự đoán hữu ích.5 Điểm quản lý nổi bật là bạn cần phải thực hiện một sự đánh đổi: sự dự đoán càng chính xác sẽ yêu cầu nhiều đơn vị hơn để nghiên cứu, và để có được những đơn vị bổ sung này có thể sẽ rất tốn kém.

Cardiogram yêu cầu tần suất cao của việc thu nhập dữ liệu. Công nghệ của họ sử dụng Apple Watch để thu nhập dữ liệu trên nền tảng từng giây. Họ cần tần suất cao như vậy vì nhịp tim dao động trong ngày và sự đo lường chính xác yêu cầu sự đánh giá lặp đi lặp lại để xem liệu tỷ lệ đo được có phải là giá trị đúng với người họ đang làm nghiên cứu không. Để hoạt động, thuật toán của Cardiogram sử dụng dòng đo lường ổn định mà một thiết bị đeo tay có thể cung cấp, thay vì sự đo lường chỉ có thể làm được khi bệnh nhân ở phòng khám của bác sĩ.

Thu thập loại dữ liệu này là một sự đầu tư tốn kém. Bệnh nhân phải đeo thiết bị mọi lúc nên nó ảnh hưởng đến hoạt động hằng ngày của họ (đặc biệt là với những người không có Apple Watch). Bởi vì nó liên quan đến dữ liệu sức

khoẻ, nhiều sự lo ngại về vấn đề quyền riêng tư đã nảy sinh, vì vậy Cardiogram đã phát triển hệ thống để cải thiện quyền riêng tư nhưng với chi phí phát triển gia tăng và làm giảm khả năng của máy để cải thiện sự dự đoán từ phản hồi. Nó thu thập dữ liệu sử dụng để dự đoán thông qua ứng dụng; dữ liệu vẫn lưu lại trên máy.

Tiếp đến, chúng tôi sẽ thảo luận về điểm khác biệt giữa suy nghĩ về mặt thống kê và suy nghĩ về mặt kinh tế liên quan đến số lượng dữ liệu thu thập được. (Chúng tôi sẽ xem xét những vấn đề liên quan đến quyền riêng tư khi bàn về chiến lược ở phần thứ tư).

Quy mô kinh tế

Nhiều dữ liệu cải thiện sự dự đoán. Nhưng bạn cần bao nhiêu dữ liệu? Lợi ích của việc có thêm thông tin (cho dù là về số lượng đơn vị, loại biến hay tần suất) có thể sẽ làm tăng hoặc giảm với số lượng dữ liệu hiện có. Dưới góc nhìn của chuyên gia kinh tế, dữ liệu có thể tăng hoặc giảm theo hiệu suất quy mô.

Từ quan điểm thống kê đơn thuần, dữ liệu đã giảm theo hiệu suất quy mô. Bạn có nhiều thông tin hữu ích từ lần quan sát thứ 3 hơn là lần quan sát thứ 100 và bạn học hỏi nhiều hơn từ lần thứ 100 hơn là lần thứ 1 triệu. Khi bạn bổ sung các lần quan sát vào dữ liệu đào tạo của mình, nó sẽ trở nên ít hữu ích hơn trong việc cải thiện sự dự đoán của bạn.

Mỗi quan sát là một sự bổ sung dữ liệu cho sự dự đoán của bạn. Trong trường hợp của Cardiogram, sự quan sát là thời gian giữa những nhịp tim được ghi lại. Khi chúng tôi nói "dữ liệu đã giảm theo hiệu suất quy mô", chúng tôi muốn nói rằng nhịp tim thứ 100 đầu tiên sẽ cho bạn biết liệu người đó có nhịp tim bất thường hay không. Mỗi nhịp tim sau đó sẽ ít quan trọng hơn những nhịp tim trước đó trong việc cải thiện sự dự đoán.

Hãy nghĩ đến thời gian mà bạn cần rời đi khi bạn định đến sân bay. Nếu bạn chưa từng đến sân bay bao giờ, lần đầu tiên bạn đi sẽ đem lại nhiều thông tin hữu ích. Lần thứ hai và lần thứ ba cũng sẽ cho bạn cảm nhận về việc mất bao lâu. Tuy nhiên, cho đến lần thứ 100, có thể bạn sẽ không học hỏi được nhiều nữa. Như vậy, dữ liệu đã giảm theo hiệu suất quy mô: khi bạn càng có nhiều dữ liệu, mỗi thông tin thêm vào sau đó càng ít có giá trị hơn.

Điều này có thể không đúng từ quan điểm kinh tế, vì nó không liên quan đến việc dữ liệu có thể cải thiện sự dự đoán ra sao. Nó liên quan đến việc dữ liệu có thể cải thiện giá trị bạn nhận được từ sự dự đoán. Đôi khi sự dự đoán và kết quả đầu ra đi cùng nhau, nên sự giảm theo hiệu suất quy mô trong quan sát thống kê ngụ ý đến kết quả đầu ra mà bạn quan tâm. Tuy nhiên, mọi chuyện không phải lúc nào cũng vậy.

Người tiêu dùng có thể lựa chọn sản phẩm của bạn hoặc đối thủ của bạn. Họ có thể chỉ sử dụng sản phẩm của bạn nếu như nó gần như luôn tốt bằng hoặc hơn sản phẩm từ đối thủ của bạn. Trong nhiều trường hợp, tất cả đối thủ cạnh tranh sẽ tốt ngang nhau trong những tình huống với những dữ liệu có sẵn. Ví dụ, hầu hết các công cụ tìm kiếm cung cấp những kết quả tương tự cho những tìm kiếm phổ biến. Cho dù bạn sử dụng Google hay Bing, những kết quả từ tìm kiếm "Justin Bieber" đều tương tự nhau. Giá trị của công cụ tìm kiếm được thúc đẩy bởi khả năng cung cấp kết quả tốt hơn cho những tìm kiếm không phổ biến bằng. Hãy thử gõ từ "sự gián đoạn" (disruption) vào Google và Bing. Vào thời điểm viết cuốn sách này, Google cho ra cả hai định nghĩa trong từ điển và những kết quả liên quan đến ý tưởng của Clay Christensen về siêu đổi mới. Chín kết quả đầu tiên của Bing cho ra định nghĩa trong từ điển. Lý do chính khiến kết quả tìm kiếm của Google tốt hơn là bởi để tìm ra điều mà người dùng cần trong kết quả tìm kiếm không phổ biến đòi hỏi nhiều dữ liệu. Đa số mọi người sử dụng Google cho những tìm kiếm không phổ biến và tìm kiếm phổ biến. Việc có khả năng tìm kiếm dù chỉ là tốt hơn một chút có thể dẫn đến sự cách biệt to lớn về thị phần và doanh thu.

Vì vậy, khi dữ liệu về mặt kỹ thuật giảm theo hiệu suất quy mô – kết quả tìm kiếm thứ một tỷ sẽ ít hữu ích cho việc cải thiện công cụ tìm kiếm hơn kết quả tìm kiếm đầu tiên – từ quan điểm kinh doanh, dữ liệu có thể sẽ có giá trị nhất nếu bạn có nhiều dữ liệu tốt hơn so với đối thủ cạnh tranh. Một vài người lập luận rằng nhiều dữ liệu độc đáo sẽ lại những lợi ích không cân xứng trên thị trường.6 Sự gia tăng lượng dữ liệu mang đến những lợi ích không cân xứng trên thị trường. Do vậy, từ quan điểm kinh tế, dữ liệu trong những trường hợp đó đã tăng theo hiệu suất quy mô.

NHỮNG ĐIỂM CHÍNH

• Máy dự đoán tận dụng ba loại dữ liệu: (1) dữ liệu đào tạo để đào tạo AI, (2) dữ liệu đầu vào để dự đoán và (3) dữ liệu phản hồi để cải thiện độ chính xác

của sự dự đoán.

• Việc thu thập dữ liệu là tốn kém; nhưng đó là một khoản đầu tư. Chi phí của việc thu thập dữ liệu phụ thuộc vào lượng bạn cần và mức độ xâm nhập của quá trình thu thập như thế nào. Việc cân bằng chi phí thu thập dữ liệu với lợi ích của việc nâng cao độ chính xác của sự dự đoán là vô cùng quan trọng. Việc xác định cách tiếp cận tốt nhất yêu cầu ước tính ROI của mỗi loại dữ liệu: chi phí để có được nó là bao nhiêu và mức độ giá trị gia tăng liên quan đến độ chính xác của sự dự đoán sẽ như thế nào?

• Những lý giải về mặt thống kê và về mặt kinh tế định hình việc liệu có thêm nhiều dữ liệu sẽ mang lại giá trị hơn hay không. Từ quan điểm thống kê, dữ liệu đã giảm theo hiệu suất quy mô. Mỗi đơn vị dữ liệu bổ sung cải thiện sự dự đoán ít hơn những dữ liệu trước đó; lần quan sát thứ 10 cải thiện sự dự đoán nhiều hơn so với lần quan sát thứ 100. Về mặt kinh tế, mối quan hệ này là mơ hồ. Việc thêm nhiều dữ liệu vào khối lượng lớn dữ liệu hiện có sẽ tốt hơn việc thêm vào khối lượng nhỏ - ví dụ, nếu lượng dữ liệu bổ sung cho phép hiệu suất của máy dự đoán vượt ngưỡng từ không thể sử dụng thành sử dụng được, hoặc từ dưới mức hiệu suất quy định thành vượt mức, hoặc từ việc kém hơn đối thủ cạnh tranh thành hơn đối thủ cạnh tranh. Do vậy, các tổ chức cần hiểu mối quan hệ giữa việc bổ sung thêm dữ liệu, nâng cao độ chính xác của sự dự đoán và gia tăng giá trị tạo thành.

6Sự phân chia lao động mới

M

ỗi lần bạn thay đổi một tài liệu điện tử, những sự thay đổi đó đều có thể được lưu lại. Với đa số chúng ta, nó không hẳn là cách hữu ích để theo dõi những thay đổi, nhưng với Ron Glozman, đây là cơ hội sử dụng AI trong dữ liệu để dự đoán những sự thay đổi. Vào năm 2015, Glozman sáng lập một công ty khởi nghiệp tên là Chisel, công ty đầu tiên có sản phẩm sở hữu tài liệu pháp lý và dự đoán những thông tin nào là bảo mật. Sản phẩm này có giá trị với những công ty luật bởi khi họ được yêu cầu tiết lộ tài liệu, họ phải biên soạn lại những thông tin bảo mật. Trong lịch sử, việc biên soạn được thực hiện bởi con người bằng cách đọc lại tài liệu và biên soạn lại thông tin bảo mật. Cách tiếp cận của Glozman hứa hẹn sẽ giúp tiết kiệm thời gian và công sức.

Máy biên soạn làm việc hiệu quả, nhưng không hoàn hảo. Đôi lúc máy chỉnh sửa sai thông tin cần tiết lộ hoặc máy không tìm được thông tin cần bảo mật. Để đạt được chuẩn mực pháp lý, cần phải có sự giúp đỡ của con người. Trong giai đoạn thử nghiệm, máy của Chisel đưa ra gợi ý thông tin gì cần biên soạn lại và con người từ chối hoặc chấp nhận sự gợi ý đó. Như vậy, khi kết hợp làm việc đồng nghĩa với việc tiết kiệm nhiều thời gian, trong khi đạt được tỉ lệ sai sót thấp hơn so với khi con người tự làm việc. Sự phân chia lao động con người-máy móc này hoạt động hiệu quả vì nó vượt qua được những điểm yếu của con người về tốc độ và sự chú ý cũng như những điểm yếu của máy móc về việc diễn giải văn bản.

Con người và máy móc đều có những thất bại. Nếu không biết chúng là gì thì chúng ta không thể đánh giá cách mà máy móc và con người nên làm việc cùng nhau để tạo ra sự dự đoán. Vì sao? Bởi vì một tư duy kinh tế của Adam Smith từ thế kỷ 18 về sự phân chia lao động bao gồm sự phân chia vai trò dựa trên những điểm mạnh tương đối. Ở đây, sự phân chia lao động là giữa con người và máy móc trong việc tạo ra sự dự đoán. Vậy sự phân chia lao động được hiểu là chúng ta sẽ phải xác định những yếu tố dự đoán nào con người hay máy móc thực hiện tốt nhất. Điều này cho phép chúng ta xác định được những vai trò cụ thể.

Việc mà con người dự đoán không giỏi

Một bài thí nghiệm tâm lý trước đây cho các đối tượng một chuỗi X và O ngẫu nhiên và yêu cầu con người dự đoán chữ cái tiếp theo sẽ là gì. Ví dụ, họ có thể thấy chuỗi:

O X X O X O X O X O X X O O X X O X O X X X O X X

Với một chuỗi như vậy, đa số mọi người nhận ra rằng có nhiều X hơn O một chút – nếu bạn đếm, bạn sẽ thấy có 60% X, 40% O – vì vậy đa số họ đoán X, nhưng đưa vào một số O để thể hiện sự cân bằng. Tuy nhiên, nếu bạn muốn tối đa hóa cơ hội dự đoán đúng, bạn sẽ luôn chọn X. Vậy thì 60% bạn sẽ đúng. Nếu bạn ngẫu nhiên ở tỉ lệ 60/40, như đa số người tham gia vào thí nghiệm, sự dự đoán của bạn sẽ đúng 52%, chỉ tốt hơn một chút so với nếu bạn không quan tâm đến việc đánh giá tần suất tương đối của X và O và thay vào đó là đoán chữ cái này hoặc chữ cái kia (50/50).[1]

Những thí nghiệm như vậy nói cho chúng ta biết con người không giỏi trong việc thống kê, ngay cả trong những tình huống mà họ không đến nỗi quá tệ ở việc đánh giá khả năng. Máy dự đoán sẽ không mắc lỗi như vậy. Nhưng có lẽ con người không quá để tâm đến những việc như vậy, vì họ cảm thấy rằng mình chỉ đang chơi một trò chơi. Liệu họ có mắc những lỗi tương tự nếu kết quả không giống như một trò chơi? Câu trả lời – được chứng minh qua nhiều thí nghiệm bởi hai nhà tâm lý học Daniel Kahneman và Amos Tversky – là chắc chắn có.[2]

Khi hai nhà tâm lý quan sát những người tham gia nghiên cứu xem xét hai bệnh viện – một bệnh viện với 45 ca sinh mỗi ngày và bệnh viện còn lại với 15 ca sinh mỗi ngày – và hỏi rằng bệnh viện nào sẽ có nhiều ngày mà số bé trai được sinh ra chiếm 60% trở lên hơn, rất ít người trả lời đúng – bệnh viện có quy mô nhỏ hơn. Bệnh viện có quy mô nhỏ hơn là đáp án đúng vì số lượng sự kiện nhiều hơn (ở trường hợp này, ca sinh), khả năng cao rằng kết quả mỗi ngày sẽ gần với mức trung bình hơn (ở trường hợp này, 50%). Do đó, bệnh viện có quy mô nhỏ hơn – là đáp án đúng vì họ có ít ca sinh hơn – có khả năng đem lại kết quả vượt xa mức trung bình.

Nhiều cuốn sách đã viết về những phương pháp tiếp cận bằng cảm tính và

thành kiến như vậy.[3] Nhiều người thấy khó khăn với việc đưa ra dự đoán dựa trên nguyên tắc thống kê hợp lý, đó chính xác là lý do vì sao họ cho mời những chuyên gia. Thật không may, những chuyên gia đó có thể thể hiện những thành kiến và khó khăn tương tự với số liệu thống kê khi đưa ra quyết định. Những thành kiến như vậy gây ảnh hưởng xấu đến những ngành đa dạng như y học, luật, thể thao và kinh doanh. Tversky, cùng với những chuyên gia nghiên cứu tại Trường Y Harvard, giới thiệu đến bác sĩ hai phương pháp điều trị ung thư phổi: xạ trị hoặc phẫu thuật. Tỉ lệ sống sót ghi nhận trong năm năm khuyên nên phẫu thuật. Hai nhóm người tham gia đã nhận được nhiều cách biểu đạt thông tin khác nhau về tỉ lệ thành công ngắn hạn của phẫu thuật, phương pháp nguy hiểm hơn là xạ trị. Khi được nói rằng "tỉ lệ sống sót trong tháng đầu tiên là 90%", 84% bác sĩ chọn phẫu thuật, nhưng tỉ lệ này giảm xuống còn 50% khi được nói là "có 10% tỉ lệ tử vong trong tháng đầu tiên". Cả hai cụm đều truyền tải cùng một điều, nhưng cách các chuyên gia nghiên cứu xây dựng thông tin ảnh hưởng dẫn đến thay đổi lớn trong quyết định. Máy móc sẽ không cho ra kết quả này.

Kahneman xác định được những tình huống khác khi những chuyên gia không giỏi dự đoán phải đối mặt với thông tin phức tạp. Những bác sĩ X-quang giàu kinh nghiệm tự mâu thuẫn với bản thân họ 1/5 lần khi đánh giá kết quả X-quang. Kiểm toán viên, nhà bệnh lý học, nhà tâm lý học và quản lý cũng thể hiện sự dao động tương tự. Kahneman kết luận rằng nếu như có cách dự đoán sử dụng một công thức thay vì một con người, công thức sẽ được xem xét nghiêm túc.

Sự dự đoán không tốt từ chuyên gia là trọng tâm của bộ phim Moneyball (tạm dịch: Cuộc chiến sân cỏ) của Michael Lewis.[4] Đội bóng chày Oakland Athletics gặp phải một vấn đề sau khi ba trong số những cầu thủ giỏi nhất rời đội, đội không còn đủ nguồn tài chính để tuyển thêm cầu thủ thay thế. Giám đốc điều hành câu lạc bộ Oakland Athletics, Billy Beane (do Brad Pitt thủ vai trong phim) sử dụng một hệ thống thống kê được phát triển bởi Bill James để dự đoán hiệu suất của cầu thủ. Với hệ thống "thuật toán bóng chày" này, Beane và những chuyên gia phân tích đã bác bỏ những lời đề xuất của những người huấn luyện viên của câu lạc bộ và chọn đội riêng của họ. Cho dù với ngân sách khiêm tốn, đội đã vượt qua những đối thủ để đến với World Series vào năm 2002. Trọng tâm của cách tiếp cận mới này là việc tránh xa những chỉ số mà họ từng nghĩ là quan trọng (ví dụ như số lần cướp gôn và mức đánh

bóng trung bình) để tập trung vào những chỉ số khác (ví dụ như hiệu suất cầu thủ tại gôn và phần trăm đánh bóng trong sân). Đồng thời cũng tránh xa phương pháp tiếp cận suy nghiệm kỳ quái của vị huấn luyện viên. Với thuật toán liên quan đến quá trình đưa ra quyết định như vậy, không ngạc nhiên khi sự dự đoán được thúc đẩy bởi dữ liệu thường có khả năng vượt xa những sự dự đoán của con người trong bóng chày.

Số liệu mới được chú ý này lý giải sự đóng góp của các cầu thủ đến hiệu suất tổng thể của đội. Máy dự đoán mới cho phép đội Oakland Athletics xác định những cầu thủ chơi kém hơn so với những cầu thủ được đánh giá theo cách truyền thống và do đó có giá trị tốt hơn về mặt chi phí tương quan với sự ảnh hưởng của họ lên hiệu suất của đội. Đội Oakland Athletics đã tập trung vào những ưu tiên đó.[5]

Có lẽ thể hiện rõ ràng nhất của sự khó khăn trong dự đoán của con người, ngay cả với những chuyên gia kinh nghiệm và có ảnh hưởng lớn, đến từ nghiên cứu về những thẩm phán đưa ra quyết định về việc tại ngoại ở Hoa Kỳ.[6] Ở Hoa Kỳ, có khoảng 10 triệu quyết định như vậy mỗi năm, và việc ai đó được tại ngoại hay không là rất quan trọng đối với gia đình, công việc và những vấn đề cá nhân khác, chưa kể đến chi phí nhà tù của chính phủ. Thẩm phán cần đưa ra quyết định dựa vào việc liệu bị cáo có bỏ trốn hoặc thực hiện những tội ác khác hay không nếu được tại ngoại, chứ không phải liệu một bản án kết tội khác có khả năng xảy ra hay không. Các tiêu chí quyết định rõ ràng và được định nghĩa cụ thể. Các nhà nghiên cứu sử dụng máy tự học để phát triển thuật toán dự đoán khả năng một bị cáo cụ thể có tái phạm pháp hay bỏ trốn trong khi tại ngoại hay không. Dữ liệu đào tạo có tính bao quát: 3/4 trong số hàng triệu người được phép tại ngoại ở thành phố New York từ năm 2008 đến 2013 đều tái phạm pháp. Thông tin bao gồm những bản ráp trước đó, những tội danh bị buộc tội và thông tin về nhân khẩu học.

Máy đưa ra dự đoán tốt hơn thẩm phán. Ví dụ, với 1% số bị cáo máy phân loại là nguy hiểm nhất, nó dự đoán rằng 62% trong số đó sẽ tái phạm pháp trong thời gian tại ngoại. Tuy nhiên, thẩm phán (người không được tiếp cận với sự dự đoán của máy) có khả năng sẽ cho gần một nửa số người đó tại ngoại. Sự dự đoán của máy là chính xác một cách hợp lý, với 63% số bị cáo được phân loại là nguy hiểm thực tế đã phạm pháp trong thời gian tại ngoại và hơn một nửa số đó không xuất hiện ở ngày hầu tòa tiếp theo. 5% trong số

đó máy xác định là nguy hiểm cao đã phạm tội hiếp dâm hoặc giết người trong thời gian tại ngoại. Bằng cách làm theo những lời đề xuất của máy, thẩm phán có thể cho tại ngoại với số lượng tương đương và giảm thiểu tỉ lệ phạm tội của những người tại ngoại đến 3/4. Hoặc họ có thể giữ nguyên tỉ lệ phạm tội và bắt tù thêm một nửa số bị cáo khác.[7]

Vậy điều gì đang diễn ra ở đây? Vì sao thẩm phán lại có sự đánh giá khác biệt đến vậy so với máy dự đoán? Một khả năng có thể là thẩm phán sử dụng thông tin không có sẵn với thuật toán, ví dụ như biểu hiện và thái độ của bị cáo trên tòa án. Thông tin đó có thể hữu ích – hoặc có thể gây đánh lừa. Với tỉ lệ tái phạm tội cao như vậy của những người được tại ngoại, không vô lý khi kết luận rằng có nhiều khả năng sự dự đoán của các thẩm phán thực sự khá tệ. Nghiên cứu cung cấp nhiều bằng chứng khác để ủng hộ cho kết luận không may này.

Sự dự đoán trở nên khó khăn với con người trong tình huống này bởi sự phức tạp của những yếu tố có thể giải thích cho tỉ lệ phạm tội. Máy dự đoán tốt hơn con người ở việc phân loại các yếu tố tương tác phức tạp giữa những chỉ số khác nhau. Vậy nên, khi bạn nghĩ rằng hồ sơ tội phạm trong quá khứ có thể khiến bị cáo có khả năng bỏ chạy cao hơn, máy đã có thể phát hiện rằng đó chỉ là trong trường hợp bị cáo đã bị thất nghiệp trong một khoảng thời gian nhất định. Nói cách khác, hiệu ứng tương tác có thể là yếu tố quan trọng nhất, và khi số lượng của những tương tác đó gia tăng, khả năng con người có thể đưa ra sự dự đoán chính xác suy giảm.

Những thành kiến này không những chỉ có trong trong y học, bóng chày và luật pháp; chúng còn là yếu tố trong công việc chuyên môn. Các chuyên gia kinh tế học đã nhận ra rằng những người quản lý và người làm việc thường đưa ra dự đoán – với sự tự tin – mà không nhận ra rằng họ đang làm không tốt. Trong nghiên cứu về việc tuyển dụng của 15 công ty luật có tay nghề thấp, Mitchell Hoffman, Lisa Kahn và Danielle Li nhận ra rằng khi công ty sử dụng bài kiểm tra khách quan và có thể đánh giá trong những buổi phỏng vấn bình thường, thời gian gắn bó của ứng viên tăng 15% so với khi họ đưa ra quyết định tuyển dụng dựa vào việc phỏng vấn.[8] Với những công việc như vậy, những nhà quản lý thường được hướng dẫn để tối đa hóa thời gian gắn bó của ứng viên. Bản thân bài kiểm tra đã mang tính bao quát rộng, bao gồm khả năng nhận thức và những yếu tố thích hợp với công việc. Ngoài ra, ngay

cả khi quyết định tuyển dụng của người quản lý bị hạn chế - ngăn cản người quản lý phủ nhận kết quả kiểm tra khi điểm số không như mong muốn – tỉ lệ ứng viên gắn bó gia tăng và tỉ lệ thôi việc giảm vẫn xảy ra.

Việc mà máy dự đoán không giỏi

Cựu thư ký của Bộ Quốc phòng Donald Rumsfeld từng nói:

Có những điều đã biết là đã biết; có những điều chỉ có chúng ta biết là chúng ta biết. Chúng ta cũng biết rằng có những điều đã biết là chưa biết; tức là có những điều chúng ta biết là chưa biết. Nhưng cũng có những điều chưa biết là chưa biết – những điều mà chúng ta chưa biết là chúng ta chưa biết. Và nếu ai đó nhìn vào lịch sử của đất nước chúng ta và những quốc gia độc lập khác, họ sẽ có xu hướng thấy những quốc gia độc lập khác khó đoán hơn.[9]

Điều này cung cấp một cấu trúc hữu ích để hiểu về những điều kiện khiến máy dự đoán lưỡng lự. Đầu tiên, những điều chúng ta biết là đã biết là khi chúng ta có lượng dữ liệu phong phú, chúng ta biết rằng mình có thể đưa ra sự dự đoán tốt. Thứ hai, những điều chúng ta biết là chưa biết là khi có quá ít dữ liệu, chúng ta biết rằng việc dự đoán sẽ khó khăn. Thứ ba, những điều chúng ta chưa biết là chúng ta chưa biết là những sự kiện chưa được thu thập thông tin từ quá khứ hoặc đã có sẵn ở dữ liệu hiện tại nhưng không khả thi, nên việc dự đoán là khó khăn, mặc dù chúng ta có thể không nhận ra điều đó. Cuối cùng, một mục mà Rumsfeld không nhận ra, những điều chưa biết là đã biết, là khi một sự liên quan dường như rõ ràng trong quá khứ là kết quả của một yếu tố chưa biết hoặc chưa được quan sát nhưng có thể thay đổi theo thời gian và khiến sự dự đoán của chúng ta trở nên không đáng tin cậy.

Những điều biết là đã biết

Với khối lượng dữ liệu phong phú, máy dự đoán có thể hoạt động tốt. Máy biết rõ tình hình và nó có thể cung cấp sự dự đoán tốt. Và chúng ta biết rằng sự dự đoán sẽ tốt. Đây chính là điểm mạnh của thế hệ trí tuệ nhân tạo hiện giờ. Phát hiện lừa đảo, chẩn đoán y khoa, cầu thủ bóng chày và quyết định tại ngoại đều nằm trong mục này.

Những điều đã biết là chưa biết

Ngay cả những mô hình dự đoán tốt nhất hiện nay (và trong tương lai gần) đòi hỏi lượng dữ liệu lớn, đồng nghĩa với việc chúng ta biết rằng sự dự đoán sẽ có thể tương đối không tốt trong những tình huống mà chúng ta không có nhiều dữ liệu. Chúng ta biết rằng là chúng ta không biết: những điều đã biết là chưa biết.

Chúng ta có thể không có nhiều dữ liệu vì một số sự kiện rất hiếm xảy ra, nên việc dự đoán chúng là một thử thách. Những cuộc bầu cử tổng thống Hoa Kỳ diễn ra bốn năm một lần, những ứng viên và môi trường bầu cử thay đổi. Dự đoán kết quả bầu cử tổng thống trong một vài năm tới là điều bất khả thi. Cuộc bầu cử năm 2016 đã cho thấy rằng dự đoán kết quả trước một vài ngày, hay thậm chí là trong ngày bầu cử là rất khó. Những trận động đất lớn thường (và thật may mắn) rất hiếm nên việc dự đoán thời gian, địa điểm và độ lớn của chúng đến nay vẫn rất khó. (Các nhà địa chấn học vẫn đang nghiên cứu vấn đề này).[10]

Trái ngược với máy móc, con người đôi khi giỏi dự đoán với lượng ít dữ liệu. Chúng ta có thể nhận ra khuôn mặt chỉ sau khi thấy một hoặc hai lần, thậm chí ngay cả khi chúng ta nhìn từ góc độ khác. Chúng ta có thể nhận ra bạn học cùng lớp 4 vào 40 năm sau, cho dù có nhiều thay đổi về ngoại hình. Chúng ta cũng giỏi trong việc so sánh, đem những tình huống mới và xác định những hoàn cảnh khác có điểm tương tự đủ để hữu ích trong một môi trường mới.[11]

Các nhà khoa học máy tính đang nghiên cứu việc giảm nhu cầu dữ liệu mà máy cần, phát triển các kỹ thuật ví dụ như “học một lần” mà máy học cách dự đoán một đối tượng đủ tốt chỉ sau một lần nhìn thấy, hiện tại thì máy dự đoán chưa đủ tốt để làm điều đó.[12] Bởi vì có những điều đã biết là chưa biết và vì con người vẫn giỏi trong việc đưa ra quyết định với những điều đã biết là chưa biết, những con người quản lý máy biết rằng những tình huống như vậy có thể phát sinh và vì vậy họ có thể lập trình máy để gọi con người khi cần giúp đỡ.

Những điều chưa biết là chưa biết

Để dự đoán, một người cần cho máy biết điều gì đáng để dự đoán. Nếu một điều chưa từng xảy ra trước đó, máy không thể dự đoán được (ít nhất là nếu

không có sự phán đoán cẩn thận của con người để cung cấp sự so sánh hữu ích cho phép máy dự đoán sử dụng thông tin về một điều khác).

Nassim Nicholas Taleb tập trung nhấn mạnh về những điều chưa biết là chưa biết trong cuốn sách The Black Swan (tạm dịch: Thiên nga đen) của ông.[13]Ông nhấn mạnh rằng chúng ta không thể dự đoán những sự kiện hoàn toàn mới từ dữ liệu trong quá khứ được. Tiêu đề của cuốn sách đề cập đến sự phát hiện của người châu Âu về một loại thiên nga mới ở Úc. Với những người châu Âu của thế kỷ 18, thiên nga có màu trắng. Cho đến khi tới Úc, họ nhìn thấy một điều hoàn toàn mới và không thể đoán trước được: thiên nga đen. Họ chưa từng nhìn thấy thiên nga đen và vì vậy họ không hề có thông tin để có thể dự đoán sự tồn tại của loài thiên nga này.[14] Taleb cho rằng sự xuất hiện của những điều chưa biết là chưa biết khác có hệ quả quan trọng – không giống như sự xuất hiện của những con thiên nga đen, một sự xuất hiện ít có ảnh hưởng đến định hướng của xã hội châu Âu hay châu Úc.

Ví dụ, những năm 1990 là thời điểm tốt để gia nhập vào ngành công nghiệp âm nhạc.[15] Doanh số bán đĩa CD gia tăng và doanh thu tăng trưởng ổn định. Tương lai trông rất xán lạn. Và rồi vào năm 1999, Shawn Fanning, 18 tuổi khi đó phát triển Napster, một chương trình cho phép mọi người chia sẻ file âm nhạc miễn phí trên Internet. Nhanh chóng, mọi người bắt đầu tải xuống hàng triệu file như vậy và nền công nghiệp âm nhạc bắt đầu đi xuống. Ngành công nghiệp này hiện giờ vẫn chưa thể phục hồi.

Phải thừa nhận rằng, con người cũng tương đối không giỏi trong việc dự đoán những điều chưa biết là chưa biết. Khi đối mặt với những điều chưa biết là chưa biết, cả con người và máy móc đều thất bại.

Những điều chưa biết là đã biết

Có lẽ điểm yếu lớn nhất của máy dự đoán là chúng đôi khi cung cấp những câu trả lời sai mà chúng lại tự tin rằng mình đúng. Như chúng tôi đã mô tả ở trên, trong trường hợp những điều đã biết là chưa biết, con người hiểu rõ độ thiếu chính xác của sự dự đoán. Sự dự đoán mang đến một sự tự tin mà cho thấy sự thiếu chính xác của nó. Trong trường hợp những điều chưa biết là chưa biết, con người không có bất kỳ câu trả lời nào. Ngược lại, với những điều chưa biết là đã biết, máy dự đoán dường như có thể cung cấp một câu trả

lời vô cùng chính xác, nhưng câu trả lời đó có thể sai. Tại sao điều này lại xảy ra? Bởi vì, trong khi dữ liệu tạo thành những quyết định, dữ liệu cũng có thể đến từ những quyết định. Nếu máy không hiểu quá trình đưa ra quyết định tạo ra dữ liệu, sự dự đoán của nó có thể thất bại. Ví dụ, nếu bạn quan tâm đến việc dự đoán rằng liệu bạn sẽ sử dụng máy dự đoán trong tổ chức của bạn hay không. Bạn đang bắt đầu rất vì việc bạn đọc cuốn sách này tức là bạn đã có được yếu tố dự đoán tuyệt vời để trở thành một người quản lý có thể sử dụng máy dự đoán.

Vì sao? Vì ít nhất ba lý do sau. Đầu tiên và trực tiếp nhất, những thông tin sâu sắc trong cuốn sách này sẽ trở nên hữu ích, vì vậy việc đọc cuốn sách này sẽ khiến bạn muốn học về máy dự đoán và áp dụng những công cụ này vào trong doanh nghiệp của bạn một cách hiệu quả.

Thứ hai là lý do gọi là “nhân quả nghịch”. Bạn đang đọc cuốn sách này vì bạn đã sử dụng máy dự đoán hoặc đã có những kế hoạch nhất định để làm như vậy trong tương lai gần. Cuốn sách này không nhằm hướng tới việc sử dụng công nghệ; thay vào đó, sự sử dụng công nghệ (có thể là đang trong giai đoạn chờ) khiến bạn đọc cuốn sách này.

Thứ ba là lí do gọi là “biến bị bỏ sót”. Bạn là người quan tâm tới những xu hướng công nghệ và quản lý. Vì vậy, bạn quyết định đọc cuốn sách này. Bạn cũng sử dụng tốt những công nghệ mới như máy dự đoán trong công việc. Trong trường hợp này, những sự yêu thích tiềm tàng với công nghệ và quản lý của bạn đã khiến bạn đọc cuốn sách này và sử dụng máy dự đoán.

Đôi khi sự phân biệt ở trên không quan trọng. Nếu tất cả những gì bạn quan tâm là liệu người đọc cuốn sách này sẽ sử dụng máy dự đoán hay không, thì điều gì dẫn đến điều gì không còn quan trọng.

Đôi khi sự khác biệt đó lại quan trọng. Nếu bạn đang nghĩ về việc giới thiệu cuốn sách này cho bạn bè, bạn sẽ làm như vậy nếu cuốn sách khiến bạn trở thành một người quản lý tốt hơn với máy dự đoán. Bạn muốn biết điều gì? Bạn sẽ bắt đầu với thực tế rằng bạn đã đọc cuốn sách này. Sau đó bạn muốn nhìn thấy tương lai và quan sát xem bạn quản lý AI giỏi thế nào. Giả sử bạn nhìn thấy một tương lai toàn hảo, bạn đã cực kỳ thành công trong việc quản lý máy dự đoán, nó trở thành cốt lõi trong tổ chức của bạn và bạn cùng tổ chức thành công ngoài sức mong đợi. Vậy bạn có thể nói rằng việc đọc cuốn

sách này dẫn đến sự thành công đó?

Không. Để tìm hiểu xem việc đọc cuốn sách này có ảnh hưởng như thế nào, bạn cũng cần biết xem điều gì sẽ xảy ra nếu bạn không đọc cuốn sách này. Bạn cần quan sát điều mà những chuyên gia kinh tế và nhà thống kê học gọi là “phản thực tế”: điều gì sẽ xảy ra nếu bạn thực hiện một hành động khác. Xác định xem liệu hành động đó có tạo ra một kết quả đòi hỏi hai sự dự đoán: đầu tiên, kết quả nào sẽ xảy ra sau khi hành động đó được thực hiện, và thứ hai, kết quả nào sẽ xảy ra nếu một hành động khác được thực hiện. Nhưng điều đó là bất khả thi. Bạn sẽ không bao giờ có dữ liệu của hành động không được thực hiện.[16]

Đây là vấn đề hiện tại của sự dự đoán bằng máy. Trong cuốn sách Deep Thinking (tạm dịch: Tư duy sâu), kiện tướng cờ vua Garry Kasporov thảo luận về vấn liên quan đến thuật toán máy tự học từ thời xưa của cờ vua:

Khi Michie và một vài đồng nghiệp viết chương trình thử nghiệm cờ vua cho máy tự học vào những năm đầu 1980, nó đã cho ra một kết quả thú vị. Họ đưa hàng trăm nước đi trong hàng nghìn ván kiện tướng vào máy, hy vọng nó có thể phát hiện ra cái nào hiệu quả và cái nào không. Đầu tiên nó có vẻ như hiệu quả. Sự đánh giá nước đi trở nên chính xác hơn so với những chương trình thông thường. Vấn đề xảy ra khi họ cho nó thực sự chơi một ván cờ vua. Chương trình phát triển các phần của nó, tấn công và lập tức mất con hậu! Nó thua chỉ trong vài bước cờ, hy sinh con hậu. Tại sao nó lại làm vậy? Khi một kiện tướng cờ hy sinh con hậu của anh ta, dường như đó là một nước đi tuyệt vời và quan trọng. Đối với máy, được đào tạo bởi những ván kiện tướng cờ, hy sinh con hậu của nó rõ ràng chính là chìa khóa thành công![17]

Máy đã đảo ngược trình tự nhân quả. Không hiểu rằng việc kiện tướng hy sinh con hậu chỉ khi để có chiến thắng nhanh và gọn, máy đã học rằng nó sẽ chiến thắng nhanh chóng ngay sau khi hy sinh con hậu. Vậy nên việc hy sinh con hậu một cách sai lầm dường như là cách dẫn đến chiến thắng. Trong khi vấn đề cụ thể này về sự dự đoán của máy đã được giải quyết, sự đảo ngược nhân quả vẫn còn là một thử thách lớn cho máy dự đoán.

Vấn đề này dường như cũng xuất hiện thường xuyên trong kinh doanh. Trong nhiều ngành công nghiệp, giá thành rẻ thường gắn liền với doanh thu thấp. Ví

dụ, trong ngành công nghiệp khách sạn, giá thành thấp khi không trong mùa du lịch, giá thành cao khi nhu cầu cao và khách sạn hết phòng. Với dữ liệu đó, một dự đoán ngây ngô có thể gợi ý rằng việc tăng giá sẽ dẫn đến việc nhiều phòng được bán ra. Một người – ít nhất với một vài hiểu biết về kinh tế - sẽ hiểu rằng việc thay đổi giá thành có thể được gây ra bởi nhu cầu cao, chứ không phải điều ngược lại. Vậy nên việc tăng giá thành sẽ không tăng doanh thu. Người này sau đó có thể làm việc với máy để xác định đúng dữ liệu (ví dụ như những sự lựa chọn cá nhân về phòng khách sạn dựa vào giá thành) và những mô hình phù hợp (mà liên quan đến mùa cũng như những nhu cầu cung và cầu khác) để dự đoán doanh thu tốt hơn ở những giá thành khác nhau. Do vậy, đối với máy, đây là điều chưa biết là đã biết, nhưng với con người, với sự hiểu rõ cách giá thành được xác định, sẽ thấy đây là điều đã biết là chưa biết hoặc có thể là điều đã biết là đã biết nếu con người có thể xây dựng mô hình chính xác của việc định giá.

Vấn đề của những điều chưa biết là đã biết và suy luận nhân quả thậm chí còn quan trọng hơn khi có sự xuất hiện của những hành vi chiến lược khác. Những kết quả tìm kiếm của Google đến từ một thuật toán bí mật. Thuật toán đó phần lớn được xác định bởi máy dự đoán có khả năng dự đoán được những liên kết mà một ai đó có khả năng bấm vào. Với một người quản lý trang web, thứ hạng càng cao đồng nghĩa với việc nhiều người dùng tới trang web hơn và doanh thu nhiều hơn. Đa số những người quản lý trang web nhận ra điều này và thực hiện tối ưu hóa công cụ tìm kiếm: họ điều chỉnh trang web của họ để cố gắng cải thiện thứ hạng của họ trên những kết quả tìm kiếm của Google. Những sự điều chỉnh này thường là để đánh lừa tính riêng biệt thuật toán, và khi thời gian qua đi, công cụ tìm kiếm sẽ trở nên đầy rẫy những thông tin rác, những liên kết mà người tìm kiếm không thực sự muốn.

Máy dự đoán đã làm tốt trong khoảng thời gian ngắn về việc dự đoán những gì mọi người sẽ bấm vào. Nhưng sau vài tuần hoặc vài tháng, sẽ có những nhà quản lý trang web tìm cách thay đổi hệ thống mà Google cần để điều chỉnh mô hình dự đoán một cách đáng kể. Sự qua lại giữa công cụ tìm kiếm và web rác của công cụ tìm kiếm xảy ra bởi vì máy dự đoán có thể bị đánh lừa. Trong khi Google đã cố gắng tạo ra một hệ thống ngăn cản việc đánh lừa đó, nó cũng nhận ra những yếu điểm của việc phụ thuộc hoàn toàn vào máy dự đoán và sử dụng sự phán đoán của con người để xử lý những web rác đó.[18]Instagram cũng trải qua một cuộc chiến trường kì với thư rác, và ứng

dụng này thường xuyên cập nhật thuật toán để ngăn chặn thư rác và những thông tin mang tính xúc phạm.[19] Nhìn chung, một khi con người đã xác định được những vấn đề như vậy, chúng sẽ không còn là những điều chưa biết là đã biết. Hoặc là họ sẽ tìm giải pháp để tạo ra dự đoán tốt, để vấn đề trở thành điều đã biết là đã biết và đòi hỏi con người và máy móc làm việc cùng nhau, hoặc họ không thể tìm được giải pháp và chúng sẽ trở thành những điều đã biết là chưa biết.

Sự dự đoán của máy có ảnh hưởng cực kỳ mạnh mẽ nhưng lại có những hạn chế. Nó không hoạt động tốt với lượng dữ liệu có hạn. Những người được đào tạo tốt có thể nhận ra những hạn chế này, cho dù là bởi những sự kiện hiếm hoi hay là do những vấn đề suy luận nhân quả, và cải thiện sự dự đoán của máy. Để làm được điều này, những người đó cần hiểu về máy.

Sự kết hợp cho dự đoán tốt hơn

Đôi khi, sự kết hợp giữa con người và máy móc tạo ra những kết quả dự đoán tốt nhất vì mỗi bên bổ sung cho khuyết điểm của nhau.

Vào năm 2016, nhóm chuyên gia nghiên cứu AI của Harvard/MIT chiến thắng Camelyon Grand Challenge, một cuộc thi phát triển máy phát hiện ung thư vú di căn từ những bộ phận sinh thiết. Thuật toán học sâu của nhóm chiến thắng đã dự đoán chính xác 92.5% so với nhà nghiên cứu bệnh lý học có xác suất dự đoán đúng 96.6%. Đây tuy có vẻ là một chiến thắng cho con người, nhưng những chuyên gia tiếp tục nghiên cứu và kết hợp thuật toán của họ cùng thuật toán của nhà nghiên cứu bệnh lý học. Kết quả là độ chính xác lên tới 99.5%.[20] Tỉ lệ sai sót của con người giảm từ 3.4% xuống còn chỉ 0.5%. Tỉ lệ sai sót giảm đến 85%.

Đây là bài toán phân chia lao động kinh điển, nhưng không phải về mặt thể chất như Adam Smith đã mô tả. Thay vào đó, nó là sự phân chia lao động về mặt nhận thức mà chuyên gia kinh tế đồng thời là người tiên phong trong lĩnh vực máy tính Charles Babbage mô tả lần đầu ở thế kỷ 19: “Hiệu quả của sự phân chia lao động, cả trong quá trình cơ học và quá trình tinh thần là nó cho phép chúng ta mua cũng như áp dụng số lượng kĩ năng và kiến thức cần thiết cho nó.”[21]

Con người và máy móc đều làm tốt ở nhiều khía cạnh dự đoán khác nhau. Nhà nghiên cứu bệnh lý học nhìn chung thường đúng khi chẩn đoán bệnh ung thư. Tình huống mà con người nói rằng đó là căn bệnh ung thư nhưng lại bị nhầm lẫn là rất hiếm. Ngược lại, AI thường chính xác hơn khi nói rằng đó không phải là căn bệnh ung thư. Con người và máy mắc phải những sai sót khác nhau. Bằng việc nhận ra những khả năng khác nhau này, sự kết hợp giữa dự đoán của con người và máy sẽ vượt qua những yếu điểm này, từ đó nhanh chóng giảm tỉ lệ sai sót.

Vậy sự kết hợp này hoạt động ra sao trong môi trường doanh nghiệp? Sự dự đoán của máy có thể nâng cao năng suất dự đoán của con người thông qua hai con đường chính. Đầu tiên là bằng cách cung cấp một sự dự đoán ban đầu mà con người có thể sử dụng để kết hợp với những đánh giá của riêng họ. Thứ hai là bằng cách cung cấp ý kiến thứ hai sau một sự thật, hoặc sau một chỉ dẫn theo dõi. Với cách này, chủ doanh nghiệp có thể đảm bảo rằng con người làm việc chăm chỉ và đặt nỗ lực vào việc dự đoán. Nếu không theo dõi, con người có thể sẽ không làm việc chăm chỉ.

Một ví dụ tuyệt vời để kiểm tra sự tương tác như vậy là sự dự đoán liên quan đến khả năng thanh toán nợ của người nộp đơn xin vay. Daniel Paravisini và Antoinette Schoar tiến hành đánh giá một ngân hàng Colombia về những doanh nghiệp nhỏ nộp đơn vay sau khi có hệ thống tính điểm tín dụng mới.[22] Hệ thống tính điểm vi tính hóa thu thập nhiều thông tin về những người nộp đơn và tổng hợp lại thành một đơn vị đo lường để dự đoán rủi ro. Sau đó ủy ban cho vay bao gồm nhân viên của ngân hàng sử dụng điểm và quá trình riêng của họ để phê duyệt, từ chối, hoặc giới thiệu khoản vay tới một người quản lý trong khu vực.

Một mẫu thử nghiệm ngẫu nhiên có đối chứng, không phải quy định quản lý, xác định liệu điểm số được giới thiệu trước hay sau quyết định. Do vậy, điểm số cung cấp một điểm tốt để đánh giá sự hiệu quả của quá trình đưa ra quyết định một cách khoa học. Một nhóm nhân viên được cung cấp số điểm ngay trước khi họ gặp khách hàng. Điều này tương tự với cách kết hợp con người với máy móc đầu tiên, trong đó sự dự đoán của máy móc ảnh hưởng đến quyết định của con người. Một nhóm nhân viên khác không được cung cấp số điểm trước khi họ đưa ra sự đánh giá đầu tiên. Điều này tương tự với cách kết hợp con người với máy móc thứ hai, sự dự đoán của máy giúp theo dõi chất

lượng quyết định của con người. Sự khác biệt giữa lần thứ nhất và lần thứ hai là liệu số điểm có cung cấp thông tin để con người đưa ra quyết định hay không.

Trong cả hai trường hợp, điểm số hoạt động hiệu quả, mặc dù sự cải thiện là lớn nhất khi điểm số được cung cấp trước. Trong trường hợp đó, ủy ban đưa ra những quyết định tốt hơn và yêu cầu người quản lý giúp đỡ ít hơn. Trong trường hợp khác, khi ủy ban có số điểm sau đó, quá trình đưa ra quyết định cải thiện bởi vì sự dự đoán giúp những nhà quản lý cao cấp theo dõi những ủy ban. Nó gia tăng động lực cho ủy ban để đảm bảo chất lượng quyết định của họ.

Sự ghép cặp con người-máy dự đoán để tạo ra một sự dự đoán tốt hơn đòi hỏi sự hiểu rõ những hạn chế của con người và máy móc. Trong trường hợp của những ủy ban cho vay, con người có thể đưa ra sự dự đoán mang tính thiên vị hoặc họ có thể trốn tránh, trong khi máy móc có thể thiếu thông tin quan trọng. Trong khi chúng ta thường coi trọng tinh thần đồng đội khi con người hợp tác, chúng ta có thể không nghĩ đến sự ghép cặp con người-máy như là một đội. Để con người có thể giúp sự dự đoán của máy trở tốt hơn, hoặc ngược lại, điều quan trọng là hiểu rõ yếu điểm của con người và máy, sau đó kết hợp cả hai sao cho có thể vượt qua những yếu điểm.

Sự dự đoán ngoại lệ

Lợi ích lớn nhất của máy dự đoán là chúng có thể vượt qua con người. Một nhược điểm là chúng gặp khó khăn trong việc đưa ra sự dự đoán trong những trường hợp không phổ biến khi không có nhiều dữ liệu trong lịch sử. Kết hợp lại, điều này đồng nghĩa với việc nhiều sự kết hợp giữa con người-máy sẽ xảy ra dưới dạng “sự dự đoán ngoại lệ”.

Như chúng ta đã thảo luận, máy dự đoán học hỏi khi dữ liệu có nhiều, điều này xảy ra khi chúng đối mặt với nhiều quy luật hoặc viễn cảnh thường xuyên. Trong những tình huống này, máy dự đoán hoạt động mà không cần con người chú ý. Ngược lại, khi một yếu tố ngoại lệ phát sinh – một viễn cảnh không theo quy luật – nó nhờ đến sự giúp đỡ của con người, và sau đó con người đặt nhiều nỗ lực hơn để cải thiện và xác nhận sự dự đoán. “Sự dự đoán ngoại lệ” này chính xác là những gì xảy ra với ủy ban cho vay ở ngân hàng Colombia.

Ý tưởng về sự dự đoán ngoại lệ đã có tiền lệ trong kỹ thuật “quản lý ngoại lệ”. Khi đưa ra sự dự đoán, con người, trong nhiều khía cạnh, là người giám sát máy dự đoán. Người quản lý có nhiều công việc khó khăn; để tiết kiệm thời gian, và tạo ra mối quan hệ làm việc lý tưởng thì con người chỉ tập trung vào những gì thực sự cần thiết. Điều đó có nghĩa rằng con người có thể dễ dàng tận dụng những lợi ích của máy dự đoán trong những dự đoán theo quy luật.

Sự dự đoán ngoại lệ rất quan trọng với sản phẩm đầu tiên của Chisel. Sản phẩm đầu tiên của Chisel, điều mà chúng ta đã thảo luận ở đầu chương này, cần nhiều tài liệu khác nhau, nhiều thông tin đã được xác định và biên soạn bảo mật. Thủ tục tốn thời gian này phát sinh nhiều tình huống pháp lý như tài liệu có thể được tiết lộ cho những bên khác hoặc công khai, nhưng chỉ khi một số thông tin nhất định được ẩn đi.

Máy biên soạn của Chisel dựa vào sự dự đoán ngoại lệ thông qua việc duyệt sơ công việc đó.[23] Cụ thể, một người dùng có thể thiết lập hiệu quả máy biên soạn sao cho mạnh hay yếu. Ngưỡng bị chặn của máy biên soạn mạnh có thể cao hơn phiên bản yếu hơn. Ví dụ, nếu bạn lo lắng về việc thông tin bảo mật không được biên soạn, bạn có thể chọn mức độ mạnh. Nhưng nếu bạn lo lắng về việc tiết lộ quá ít thông tin, bạn có thể chọn mức độ yếu. Chisel cung cấp một giao diện dễ sử dụng cho người dùng để có thể kiểm tra lại những phần biên soạn và chấp nhận hoặc từ chối chúng. Nói cách khác, mỗi sự biên soạn đều là một đề xuất thay vì là một quyết định cuối cùng. Quyền hạn lớn nhất vẫn thuộc về con người.

Sản phẩm của Chisel kết hợp con người và máy để vượt qua những yếu điểm của cả hai. Máy hoạt động nhanh hơn con người và cung cấp sự đo lường nhất quán trên các tài liệu. Con người có thể can thiệp khi máy không có đủ dữ liệu để có thể đưa ra một sự dự đoán tốt hơn.

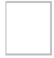
NHỮNG ĐIỂM CHÍNH

• Con người, bao gồm cả những chuyên gia chuyên nghiệp, có thể đưa ra những dự đoán không tốt dưới những điều kiện cụ thể. Con người thường chú trọng vào thông tin nổi bật và không tính đến những thuộc tính thống kê. Nhiều nghiên cứu khoa học ghi lại được những thiếu sót này trên nhiều ngành nghề. Hiện tượng này được minh họa trong bộ phim *Moneyball*.

• Máy móc và con người có những điểm mạnh và điểm yếu riêng tùy vào bối cảnh dự đoán. Khi máy dự đoán được cải thiện, các doanh nghiệp cần phải điều chỉnh, phân chia công việc giữa con người và máy móc. Máy dự đoán thường tốt hơn con người ở việc xác định những yếu tố trong những tương tác phức tạp, đặc biệt là trong những trường hợp có nhiều dữ liệu. Khi càng nhiều không gian cho những sự tương tác như vậy phát triển, khả năng con người đưa ra những dự đoán chính xác suy giảm, đặc biệt là tương quan với máy móc. Tuy nhiên, con người thường giỏi hơn máy trong việc hiểu quá trình tạo ra dữ liệu có ưu thế cho việc dự đoán, đặc biệt là trong những trường hợp ít dữ liệu. Chúng tôi mô tả sự phân loại trường hợp dự đoán (những điều đã biết là đã biết, những điều đã biết là chưa biết, những điều chưa biết là đã biết và những điều chưa biết là chưa biết) sẽ hữu ích trong việc phân chia lao động phù hợp.

• Máy dự đoán sẽ phát triển. Chi phí đơn vị cho mỗi dự đoán giảm khi tần suất tăng. Sự dự đoán của con người không phát triển theo cách như vậy. Tuy nhiên, con người có những mô hình nhận thức về cách thế giới hoạt động và do vậy có thể đưa ra dự đoán dựa vào lượng dữ liệu ít. Do vậy, chúng tôi mong chờ một sự gia tăng dự đoán ngoại lệ của con người mà ở đó máy tạo ra nhiều dự đoán nhất bởi chúng được dự đoán dựa trên quy luật, dữ liệu thông thường, nhưng khi những sự kiện hiếm có xảy ra, máy nhận ra nó không thể tự tin đưa ra dự đoán mà sẽ cần sự giúp đỡ của con người. Con người sẽ cung cấp sự dự đoán ngoại lệ.

PHẦN 2QUÁ TRÌNH ĐƯA RA QUYẾT ĐỊNH



7Mở ra những quyết định

C

húng ta thường liên hệ quá trình đưa ra quyết định với những quyết định lớn: Tôi có nên mua căn nhà này không? Tôi có nên học trường này không? Tôi có nên cưới người này không? Quả nhiên, những quyết định thay đổi cuộc đời này tuy ít nhưng rất quan trọng.

Nhưng chúng ta cũng luôn đưa ra những quyết định nhỏ: Tôi có nên tiếp tục ngồi trên chiếc ghế này không? Tôi có nên tiếp tục đi xuống con đường này không? Tôi có nên tiếp tục trả tiền hoá đơn tháng này không? Và giống như lời ca về ý chí tự do của ban nhạc rock Rush vĩ đại từ Canada: “Nếu bạn chọn không đưa quyết định, thì bạn vẫn đã đưa ra lựa chọn.” Chúng ta xử lý nhiều quyết định nhỏ một cách tự động, có lẽ bằng việc mặc định chấp nhận và lựa chọn tập trung tất cả sự chú ý vào những quyết định lớn. Tuy nhiên, quyết định không lựa chọn cũng là một quyết định.

Quá trình đưa ra quyết định là cốt lõi của hầu hết các nghề nghiệp. Giáo viên quyết định cách giáo dục học sinh của họ, những đứa trẻ có tính cách và cách học khác nhau. Những người quản lý quyết định tuyển ai vào đội của họ và ai cần được thăng chức. Những người làm tạp vụ quyết định cách xử lý những sự kiện bất ngờ như sự cố tràn nước và những điều kiện an toàn trong nguy hiểm tiềm tàng. Những người lái xe quyết định cách phản ứng với những tuyến đường bị chặn và tai nạn giao thông. Cảnh sát quyết định cách xử lý những cá nhân đáng ngờ và những tình huống có khả năng nguy hiểm. Các bác sĩ quyết định loại thuốc nào để kê đơn và khi nào cần thực hiện những bài kiểm tra đắt đỏ. Bố mẹ quyết định thời lượng xem TV phù hợp với con của họ.

Những quyết định như vậy thường xảy ra dưới những điều kiện không chắc chắn. Giáo viên không biết chắc liệu một đứa trẻ cụ thể nào đó sẽ học tốt hơn nhờ cách tiếp cận giảng dạy này hay cách tiếp cận khác. Người quản lý không biết chắc liệu một ứng viên nộp đơn xin việc có thể hiện tốt hay không. Bác sĩ không biết chắc liệu có cần phải tiến hành một xét nghiệm đắt đỏ hay không. Họ đều cần dự đoán.

Nhưng sự dự đoán không phải là một quyết định. Đưa ra quyết định đòi hỏi áp dụng sự đánh giá một dự đoán và rồi thực hiện. Trước khi có những sự tiến bộ gần đây trong trí tuệ nhân tạo, sự khác biệt này chỉ đơn thuần là sự quan tâm về mặt học thuật bởi vì con người luôn thực hiện dự đoán và đánh giá cùng lúc. Hiện giờ, những sự tiến bộ trong dự đoán bởi máy móc đồng nghĩa với việc chúng ta cần phải xác định cấu trúc của một quyết định.

Cấu trúc của một quyết định

Máy dự đoán sẽ có tác động tức thời ở giai đoạn quyết định. Nhưng sự quyết định có sáu yếu tố quan trọng khác (xem hình 7-1). Khi một người (hoặc một vật) đưa ra quyết định, họ sẽ nhận vào thông tin đầu vào từ thế giới quan cho phép sự dự đoán đó xảy ra. Sự dự đoán đó là có thể vì đã có sự đào tạo về mối quan hệ giữa các loại dữ liệu khác nhau và loại dữ liệu nào liên quan gần nhất với tình huống. Kết hợp dự đoán cùng với sự đánh giá những vấn đề quan trọng, người quyết định có thể lựa chọn *hành động*. Hành động dẫn đến *một kết quả* (thường có một thành tựu liên quan nào đó hoặc sự trả giá). Kết quả là hệ quả của quyết định. Nó là yếu tố cần thiết để có thể cung cấp một bức tranh hoàn chỉnh. Kết quả cũng có thể cung cấp phản hồi để giúp cải thiện sự dự đoán tiếp theo.



Hãy tưởng tượng bạn bị đau chân và đến gặp bác sĩ. Bác sĩ gặp bạn, chụp X-quang, xét nghiệm máu, hỏi bạn một vài câu hỏi và thu được thông tin đầu vào. Sử dụng thông tin đó, đồng thời dựa vào những năm kinh nghiệm ở trường y và những bệnh nhân có thể giống hoặc không giống bạn (đó là đào tạo và phản hồi), bác sĩ sẽ đưa ra chẩn đoán: “Bạn có khả năng là bị chuột rút cơ bắp, tuy nhiên cũng có khả năng là bạn bị tụ máu.”

Cùng với quá trình này là sự đánh giá. Đánh giá của bác sĩ sẽ dựa vào những yếu tố dữ liệu khác (bao gồm trực giác và kinh nghiệm). Giả sử nếu đó là chuột rút cơ bắp, vậy thì phương pháp điều trị là nghỉ ngơi. Nếu là tụ máu, thì phương pháp điều trị sẽ là thuốc không có tác dụng phụ lâu dài, nhưng có thể gây ra những sự khó chịu cho nhiều người. Nếu bác sĩ điều trị nhầm cơ bắp bị chuột rút của bạn thành tụ máu, vậy thì bạn sẽ cảm thấy không thoải mái trong một khoảng thời gian ngắn. Nếu bác sĩ điều trị tụ máu bằng việc nghỉ ngơi, vậy thì rất có thể sẽ có những biến chứng nguy hiểm hoặc thậm chí là tử

vong. Sự đánh giá bao gồm việc xác định những sự trả giá liên quan đến mỗi kết quả khả thi, bao gồm cả những kết quả liên quan đến những quyết định “đúng đắn” và những sai sót (trong trường hợp này, sự trả giá liên quan đến việc chữa bệnh, sự khó chịu, và những biến chứng nghiêm trọng). Xác định sự trả giá cho tất cả những kết quả khả thi là bước quan trọng trong việc quyết định khi nào nên chọn phương pháp điều trị bằng thuốc, sự khó chịu vừa phải và giảm nguy cơ biến chứng nguy hiểm, với khi nào nên chọn việc nghỉ ngơi. Vậy nên, khi đánh giá sự dự đoán, bác sĩ đưa ra quyết định, có lẽ dựa vào tuổi của bạn và ưu tiên rủi ro, bạn nên theo phương pháp điều trị chuột rút cơ bắp, ngay cả khi có khả năng thấp là bạn bị bệnh tụ máu. Cuối cùng là hành động điều trị và quan sát kết quả: Liệu cơn đau ở chân bạn có hết không? Liệu có những biến chứng khác phát sinh không? Bác sĩ có thể sử dụng kết quả quan sát là phản hồi để thông báo cho dự đoán tiếp theo.

Bằng việc chia nhỏ một quyết định thành nhiều yếu tố, chúng ta có thể đánh giá hoạt động nào của con người sẽ suy giảm giá trị và hoạt động nào sẽ tăng bởi sự dự đoán tốt hơn của máy móc. Đặc biệt, đối với bản thân sự dự đoán, máy dự đoán nhìn chung là sự thay thế tốt hơn cho sự dự đoán của con người. Khi sự dự đoán của máy càng thay thế nhiều sự dự đoán của con người, giá trị của sự dự đoán của con người sẽ suy giảm. Nhưng điểm chính là, trong khi sự dự đoán là thành phần quan trọng của bất kỳ quyết định nào, nó không phải yếu tố duy nhất. Những yếu tố khác của sự quyết định – sự đánh giá, dữ liệu và sự hành động – cho đến lúc này vẫn đang tồn tại trong tầm tay của con người. Chúng chính là sự bổ sung cho sự dự đoán, đồng nghĩa với việc chúng sẽ tăng giá trị nếu như sự dự đoán có giá thành rẻ. Ví dụ, chúng ta có thể sẽ nỗ lực hơn khi áp dụng sự đánh giá vào những quyết định mà chúng ta trước đó đã quyết định là không quyết định (ví dụ chấp nhận sự mặc định) bởi vì máy dự đoán hiện giờ cung cấp sự dự đoán tốt hơn, nhanh hơn và có giá thành rẻ hơn. Trong trường hợp đó, nhu cầu cho sự đánh giá của con người sẽ gia tăng.

Mất kiến thức

“The Knowledge” là chủ đề của một cuộc thử nghiệm mà những người lái xe London thực hiện để lái những chiếc taxi đen nổi tiếng của thành phố. Cuộc thử nghiệm bao gồm biết vị trí của hàng ngàn điểm và con đường khắp thành phố và – đây là phần khó nhất – dự đoán con đường ngắn nhất hoặc nhanh

nhất giữa hai điểm trong một thời điểm bất kỳ trong ngày. Lượng thông tin của một thành phố bình thường đã rất đáng kinh ngạc, nhưng London không phải là một thành phố bình thường. Nó là một quần thể của những ngôi làng và thị trấn độc lập cùng nhau phát triển trong suốt 2.000 năm để trở thành một đô thị toàn cầu.

Để vượt qua cuộc thử nghiệm này, những tài xế tiềm năng cần một điểm số gần như hoàn hảo. Không ngạc nhiên khi để đậu cuộc thử nghiệm này, những người lái xe mất trung bình ba năm, bao gồm thời gian không những tìm bản đồ mà còn lái xe quanh thành phố trên những xe gắn máy để ghi nhớ và hình dung. Nhưng một khi họ hoàn thành cuộc thử nghiệm này, huy chương xanh danh dự cũng được coi là một loại kiến thức.[1]

Bạn biết câu chuyện này sẽ đi đến đâu. Một thập kỷ trước, kiến thức của tài xế taxi Lonbon chính là lợi thế cạnh tranh của họ. Không ai có thể cung cấp mức độ dịch vụ tương tự. Những ai đáng lẽ sẽ đi bộ đều đi xe taxi chỉ vì những người lái xe biết đường. Nhưng chỉ năm năm sau, một thiết bị di động đơn giản gọi là GPS - hay hệ thống điều hướng vệ tinh - đã cho những người lái xe truy cập vào dữ liệu và những sự dự đoán từng là siêu năng lực của những người lái xe. Ngày nay, siêu năng lực tương tự có sẵn trên hầu hết các điện thoại. Mọi người sẽ không bị lạc. Mọi người biết tuyến đường nhanh nhất. Và giờ điện thoại thậm chí còn tốt hơn vì nó được cập nhật trong thời gian thực với thông tin giao thông.

Những người lái xe từng đầu tư ba năm nghiên cứu để học “The Knowledge” không biết rằng một ngày họ sẽ phải cạnh tranh với máy dự đoán. Qua năm tháng, họ lưu giữ hình ảnh bản đồ vào trí nhớ, những tuyến đường được kiểm tra thực nghiệm và điền vào chỗ trống những kiến thức chung của họ. Hiện giờ, những ứng dụng điều hướng có quyền truy cập vào cùng loại dữ liệu bản đồ và có thể, thông qua sự kết hợp của các thuật toán và đào tạo dự đoán, tìm được tuyến đường tốt nhất khi được yêu cầu, thông qua dữ liệu giao thông trong thời gian thực mà người lái xe khó lòng biết được.

Nhưng số phận của những người lái xe taxi không chỉ dựa vào khả năng dự đoán “The Knowledge” của các ứng dụng mà còn dựa vào những yếu tố quan trọng khác để chọn con đường tốt nhất từ điểm A đến điểm B. Đầu tiên, những người lái xe taxi có thể điều khiển xe. Thứ hai, họ có những cảm biến

- mắt và tai của họ – có thể đưa những dữ liệu ngắn gọn tới não bộ của họ để đảm bảo rằng họ sử dụng kiến thức của mình một cách tốt nhất. Nhưng những người khác cũng làm vậy. Không có tài xế taxi London nào trở nên tệ hơn trong công việc vì những ứng dụng điều hướng. Thay vào đó, hàng triệu người không phải là lái xe taxi trở nên tốt hơn nhiều. Kiến thức của người lái xe taxi không còn là nguồn hàng hiếm nữa, điều này dẫn đến những nền tảng đi chung xe như Uber.

Những người lái xe khác sử dụng "The Knowledge" với điện thoại và dự đoán tuyến đường nhanh nhất đồng nghĩa với việc họ có thể cung cấp những dịch vụ tương tự. Khi sự dự đoán chất lượng cao của máy có giá thành rẻ, sự dự đoán của con người sụt giảm giá trị, nên những người lái xe cũng gánh chịu hậu quả. Số lượt đi taxi đen ở London sụt giảm. Thay vào đó, những người còn lại cung cấp dịch vụ tương tự. Những người này còn có kĩ năng lái xe và những giác quan của con người, những tài sản bổ sung có giá trị tăng cao khi sự dự đoán có giá thành rẻ.

Đương nhiên, những chiếc xe tự lái có thể sẽ thay thế những kĩ năng và giác quan này, nhưng chúng ta sẽ trở lại với câu chuyện đó sau. Quan điểm của chúng tôi ở đây là để hiểu rõ sức ảnh hưởng từ sự dự đoán của máy đòi hỏi phải hiểu rõ nhiều khía cạnh khác nhau của quyết định, như được mô tả ở phần cấu trúc của một quyết định.

Bạn có nên cầm ô theo không?

Cho đến bây giờ, chúng ta vẫn chưa đưa ra định nghĩ thực sự chính xác về sự đánh giá. Để giải thích nó, chúng tôi giới thiệu công cụ giúp đưa ra quyết định: cây quyết định.[2] Nó đặc biệt hữu ích cho những quyết định không chắc chắn, khi bạn không chắc về việc gì sẽ xảy ra nếu bạn thực hiện một sự lựa chọn cụ thể.

Hãy xem xét một sự lựa chọn quen thuộc mà bạn có thể đối mặt. Bạn có nên cầm theo ô khi đi bộ không? Bạn có thể nghĩ rằng ô là thứ bạn che trên đầu để không dính ướt và bạn đã đúng. Nhưng chiếc ô cũng là một hình thức bảo hiểm, trong trường hợp này, chống lại khả năng mưa. Vậy nên, khuôn mẫu sau đây ứng dụng cho bất kỳ hình thức bảo hiểm nào liên quan đến quyết định để giảm thiểu rủi ro.

Rõ ràng, nếu bạn biết rằng trời sẽ không mưa, bạn sẽ để ô ở nhà. Mặt khác, nếu bạn biết là trời sẽ mưa, thì bạn sẽ chắc chắn mang nó theo cùng. Trong hình 7-2, chúng tôi thể hiện điều này bằng cách sử dụng một biểu đồ giống hình cái cây. Ở gốc cây là hai nhánh thể hiện những sự lựa chọn bạn có thể đưa ra: “mang ô” hoặc “để ô ở nhà”. Mở rộng từ đó là hai nhánh thể hiện việc bạn không chắc chắn: “mưa” với “nắng”. Nếu không có dự báo thời tiết tốt, bạn không thể biết chắc. Bạn có thể biết rằng, ở thời điểm này năm ngoái, khả năng nắng cao gấp ba lần khả năng mưa. Điều này sẽ cho bạn 1/3 khả năng nắng và 1/4 khả năng mưa. Đây là sự dự đoán của bạn. Cuối cùng, ở ngọn các nhánh là những hệ quả. Nếu bạn không mang theo ô và trời mưa, bạn sẽ bị ướt…



Bạn nên đưa ra quyết định gì? Đây là giai đoạn mà sự đánh giá bắt đầu xuất hiện. Sự đánh giá là quá trình xác định thành quả của một hành động cụ thể trong một môi trường cụ thể. Đó là về việc thực hiện được mục tiêu bạn thực sự theo đuổi. Sự đánh giá bao gồm xác định điều chúng ta gọi là “chức năng thành quả”, những thành quả tương ứng, những sự trả giá liên quan đến việc thực hiện những hành động cụ thể và cho ra những kết quả cụ thể. Bị ướt hay không? Có nên mất công mang theo ô hay không?

Hãy giả sử bạn thích không bị ướt và không muốn mang theo ô (bạn đánh giá nó 10 trên 10) hơn là việc không bị ướt mà phải mang theo ô (8 trên 10) hơn là việc bị ướt (một con số 0 tròn trĩnh). (Xem hình 7-3). Điều đó đủ khiến bạn hành động. Với sự dự đoán khả năng mưa 1/4 và đánh giá những sự trả giá cho việc bị ướt hoặc mang theo ô, bạn có thể tính toán được sự trả giá trung bình của bạn giữa việc mang và không mang ô. Dựa vào đó, tốt hơn là bạn nên mang ô (sự trả giá trung bình là 8) hơn là không mang nó (sự trả giá trung bình là 7.5).3 Nếu bạn thực sự ghét việc mang theo ô (6 trên 10), sự đánh giá của bạn về độ ưu tiên có thể được xem xét. Trong trường hợp này, sự trả giá trung bình cho việc để ô ở nhà là không đổi (7.5), trong khi sự trả giá cho việc mang theo ô lúc này là 6. Vậy nên, những người không thích ô sẽ để ô của họ ở nhà.



Ví dụ này rất đỗi bình thường: đương nhiên, ai mà không thích ô hơn cả việc bị ướt thì sẽ để ô ở nhà. Nhưng cây quyết định là một công cụ hữu ích cho việc xác định những sự trả giá cho những quyết định không hề bình thường, đó chính là cốt lõi của sự đánh giá. Ở đây, hành động là mang theo ô, sự dự đoán là mưa hay nắng, kết quả là liệu bạn có bị ướt hay không, và sự đánh giá dự đoán niềm hạnh phúc mà bạn sẽ cảm thấy ("sự trả giá") từ việc bị ướt hay không, mang hay không mang theo ô. Khi sự dự đoán trở nên tốt hơn, nhanh hơn và có giá thành rẻ hơn, chúng ta sẽ sử dụng chúng nhiều hơn để đưa ra nhiều quyết định hơn, vậy nên chúng ta cũng sẽ cần nhiều sự đánh giá của con người hơn và do vậy, giá trị của sự đánh giá từ con người sẽ tăng lên.

NHỮNG ĐIỂM CHÍNH

• Máy dự đoán rất có giá trị bởi vì (1) chúng có thể thường xuyên cho ra những sự dự đoán tốt hơn, nhanh hơn và có giá thành rẻ hơn so với con người; (2) sự dự đoán là yếu tố quan trọng trong quá trình đưa ra quyết định không chắc chắn; và (3) quá trình đưa ra quyết định phổ biến ở khắp mọi nơi trong cuộc sống kinh tế và xã hội của chúng ta. Tuy nhiên, sự dự đoán không phải là một quyết định – nó chỉ là một thành phần của sự quyết định. Những thành phần khác bao gồm sự đánh giá, hành động, kết quả và ba loại dữ liệu (dữ liệu đầu vào, dữ liệu đào tạo và dữ liệu phản hồi).

• Bằng việc chia nhỏ một quyết định ra thành nhiều thành phần như vậy, chúng ta có thể hiểu ảnh hưởng của máy dự đoán tới giá trị của con người và những tài sản khác. Tuy nhiên, giá trị của những sự bổ sung, ví dụ như những kỹ năng của con người liên quan đến thu thập dữ liệu, sự đánh giá và hành động, sẽ càng trở nên có giá trị. Trong trường hợp của những lái xe taxi London đã đầu tư ba năm để học "The Knowledge" – cách dự đoán tuyến đường nhanh nhất từ địa điểm này đến địa điểm khác ở một thời điểm bất kỳ trong ngày – không ai làm tệ hơn cả vì nhờ có máy dự đoán. Thay vào đó, rất nhiều người lái xe khác trở nên giỏi hơn trong việc chọn tuyến đường tốt nhất bằng việc sử dụng máy dự đoán. Những kỹ năng dự đoán của những người lái xe taxi không còn là một mặt hàng khan hiếm nữa.

• Sự đánh giá bao gồm việc xác định sự trả giá tương quan đối với mỗi kết quả khả thi của quyết định, bao gồm những quyết định "đúng đắn" cũng như những sai sót. Sự đánh giá đòi hỏi xác định cụ thể mục tiêu bạn đang thực sự theo đuổi và là bước quan trọng trong quá trình đưa ra quyết định. Khi máy

dự đoán đưa ra những dự đoán tốt hơn, nhanh hơn và có giá thành rẻ hơn, giá trị từ sự đánh giá của con người sẽ tăng vì chúng ta cần nhiều hơn nữa. Chúng ta có thể sẽ sẵn sàng nỗ lực và áp dụng sự đánh giá vào những quyết định mà trước đây chúng ta lựa chọn không quyết định (bằng việc chấp nhận sự mặc định).

8Giá trị của sự đánh giá

V

iệc có sự dự đoán tốt hơn làm tăng giá trị của sự đánh giá. Dù sao thì, nó cũng không biết được khả năng trời mưa nếu bạn không biết bản thân muốn khô ráo hoặc ghét việc mang theo ô đến mức nào. Máy dự đoán không cung cấp sự đánh giá. Chỉ có con người mới làm vậy, bởi vì chỉ có con người có thể diễn đạt những kết quả từ việc thực hiện những hành động khác nhau. Khi AI tiếp quản việc dự đoán, con người sẽ ít thực hiện quy trình kết hợp dự đoán-đánh giá khi đưa ra quyết định hơn và tập trung vào vai trò đánh giá. Điều này sẽ cho phép một giao diện tương tác giữa sự dự đoán của máy và sự đánh giá của con người, giống như cách bạn thực hiện những truy vấn thay thế khi tương tác với một bảng tính hoặc cơ sở dữ liệu.

Sự dự đoán tốt hơn mang đến nhiều cơ hội để cân nhắc những thành tựu của nhiều hành động khác nhau – nói cách khác, nhiều cơ hội cho sự đánh giá hơn. Và điều đó đồng nghĩa rằng sự dự đoán càng tốt hơn, nhanh hơn và có giá thành rẻ hơn sẽ cho chúng ta thực hiện nhiều quyết định hơn.

Đánh giá lừa đảo

Mạng lưới thẻ tín dụng ví dụ như Mastercard, Visa và American Express dự đoán và đánh giá mọi lúc. Họ phải dự đoán liệu những ứng viên nộp thẻ có đáp ứng đủ tiêu chuẩn của họ cho khả năng thanh toán nợ hay không. Nếu một cá nhân không thể đáp ứng, công ty sẽ từ chối thẻ tín dụng của họ. Bạn có thể nghĩ rằng đó chỉ là dự đoán đơn thuần, nhưng một yếu tố quan trọng của sự đánh giá cũng được bao gồm. Khả năng thanh toán nợ là một thang trượt, và công ty thẻ tín dụng cần phải quyết định độ rủi ro họ sẵn sàng chịu với mức lãi suất và mặc định khác nhau. Những quyết định này dẫn đến những mô hình kinh doanh khác nhau – sự khác biệt giữa thẻ bạch kim cao cấp của American Express và thẻ ở mức độ cơ bản dành cho sinh viên đại học.

Công ty cũng cần dự đoán liệu một giao dịch nào đó có hợp pháp hay không. Giống như quyết định của bạn về việc có nên mang theo ô hay không, công ty cần phải cân nhắc bốn kết quả khác nhau (hình 8-1).



Công ty cần phải dự đoán liệu một khoản phí là lừa đảo hay hợp pháp, sau đó quyết định nên cho phép hay từ chối giao dịch và rồi đánh giá mỗi kết quả (từ chối một khoản phí lừa đảo là việc làm tốt, nhưng khiến khách hàng tức giận với việc từ chối một giao dịch hợp pháp là việc làm không tốt). Nếu các công ty thẻ tín dụng hoàn hảo trong việc dự đoán lừa đảo, tất cả sẽ ổn thoả. Nhưng họ lại không như vậy. Ví dụ, công ty thẻ tín dụng của Joshua (một trong những tác giả) từ chối những giao dịch khi anh ấy mua sắm giày chạy theo định kỳ, đôi khi anh ấy mua mỗi năm một lần, thông thường là ở trung tâm mua sắm khi anh ấy đi nghỉ. Trong nhiều năm, anh ấy phải gọi đến công ty thẻ tín dụng để bỏ sự hạn chế đó.

Trộm cắp thẻ tín dụng thường xảy ra ở những trung tâm thương mại, và một vài mua sắm lừa đảo đầu tiên có thể sẽ là những thứ như giày và quần áo (dễ dàng để đổi thành tiền mặt khi trả lại hàng ở một chi nhánh khác cùng chuỗi). Và từ khi Joshua không còn thường xuyên mua quần áo và giày và ít khi đến trung tâm mua sắm, công ty thẻ tín dụng dự đoán rằng thẻ có khả năng bị đánh cắp. Đó là một sự dự đoán hợp lý.

Một vài yếu tố ảnh hưởng đến việc dự đoán liệu một chiếc thẻ có bị đánh cắp hay không bao gồm những yếu tố khái quát (loại giao dịch, ví dụ như mua giày chạy), và những yếu tố khác cụ thể hơn đối với mỗi cá nhân (trong trường hợp này, độ tuổi và tần suất). Sự kết hợp đó đồng nghĩa với việc thuật toán ngăn chặn những giao dịch sẽ rất phức tạp.

Lời hứa của AI là nó có thể đưa ra dự đoán chính xác hơn, đặc biệt trong những tình huống có sự kết hợp giữa những thông tin khái quát và thông tin cá nhân hóa. Ví dụ, nếu có dữ liệu về những năm giao dịch của Joshua (một trong những tác giả), máy dự đoán có thể học những mẫu của các giao dịch, bao gồm sự thật rằng anh ấy mua giày vào thời điểm tương tự mỗi năm. Thay vì phân loại sự mua sắm đó là sự kiện bất thường, nó có thể phân loại sự mua sắm đó là sự kiện bình thường cho một người cụ thể. Máy dự đoán có thể nhận ra những mối tương quan khác, ví dụ như mất bao nhiêu thời gian để một người mua sắm xong, tìm hiểu rằng liệu những giao dịch trong hai cửa hàng khác nhau có quá gần nhau không. Khi máy dự đoán trở nên chính xác hơn trong việc ngăn chặn những giao dịch, mạng lưới thẻ có thể trở nên tự tin

hơn khi áp đặt một hạn chế và thậm chí là liệu có nên liên lạc với khách hàng hay không. Điều này đã xảy ra. Lần mua sắm giày chạy cuối cùng của Joshua ở trung tâm mua sắm diễn ra suôn sẻ.

Nhưng cho đến khi máy dự đoán trở nên hoàn hảo trong việc dự đoán gian lận, những công ty thẻ tín dụng sẽ phải tìm ra chi phí của sai sót, điều mà đòi hỏi sự đánh giá. Giả sử sự dự đoán đó không hoàn hảo và có khả năng 10% là không chính xác. Vậy nếu những công ty từ chối giao dịch, họ đã đưa ra quyết định đúng với 90% khả năng và tiết kiệm cho mạng lưới chi phí khôi phục thanh toán liên quan đến giao dịch trái phép. Nhưng đồng thời họ cũng từ chối một giao dịch hợp pháp với 10% khả năng, để lại cho mạng lưới một lượng khách hàng không hài lòng. Để tìm ra hành động đúng đắn, họ cần có khả năng cân bằng chi phí liên quan đến phát hiện lừa đảo với chi phí liên quan đến sự không hài lòng của khách hàng. Những công ty thẻ tín dụng không tự động biết được câu trả lời đúng cho sự đánh đổi này. Họ cần tìm ra câu trả lời đó. Sự đánh giá là quá trình thực hiện việc đó. Trong trường hợp này, bởi vì khả năng giao dịch lừa đảo thường gấp chín lần so với khả năng giao dịch hợp pháp, công ty sẽ từ chối khoản phí trừ khi sự hài lòng của khác hàng quan trọng gấp chín lần so với tổn thất.

Với lừa đảo tín dụng, có rất nhiều sự trả giá có thể dễ dàng đánh giá. Khả năng cao là chi phí khôi phục có giá trị rõ ràng mà mạng lưới có thể xác định. Giả sử cho mỗi giao dịch 100 đô, chi phí khôi phục là 20 đô. Nếu chi phí cho sự không hài lòng của khách hàng ít hơn 180 đô, việc từ chối giao dịch là hợp lý (10% của 180 đô là 18 đô, bằng với 90% của 20 đô).

Lừa đảo thẻ tín dụng cũng là một quá trình đưa ra quyết định được định nghĩa rõ ràng, đó là một lý do khiến chúng tôi quay trở lại với nó, nhưng nó vẫn rất phức tạp. Ngược lại, với những quyết định khác, không chỉ có những hành động tiềm năng trở nên phức tạp hơn (không chỉ đơn thuần là sự chấp nhận hay từ chối), mà những tình huống tiềm năng (hoặc trạng thái) cũng thay đổi. Sự đánh giá đòi hỏi sự hiểu biết về kết quả của mỗi cặp hành động và tình huống. Ví dụ về thẻ tín dụng của chúng tôi chỉ có bốn kết quả (hoặc tám nếu bạn phân biệt được giữa những khách hàng có giá trị tài sản ròng lớn và những người khác). Nhưng nếu bạn có 10 hành động và 20 tình huống có thể xảy ra, vậy thì bạn sẽ phải đánh giá 200 kết quả. Khi mọi thứ trở nên phức tạp hơn, số lượng kết quả có thể trở nên áp đảo.

Các chi phí nhận thức của sự đánh giá

Những ai đã nghiên cứu về sự quyết định trong quá khứ thường cho rằng kết quả là những thứ có sẵn – chúng tồn tại. Bạn có thể thích kem sôcôla, trong khi bạn của bạn có thể thích kem gelato xoài. Việc hai người có những quan điểm khác nhau như vậy không để lại nhiều hậu quả. Tương tự, chúng tôi giả sử rằng đa số các doanh nghiệp đều tối đa hóa lợi nhuận hoặc giá trị cổ đông. Những chuyên gia kinh tế nghiên cứu lý do tại sao các công ty lựa chọn những mức giá nhất định cho sản phẩm của họ đã thấy rằng việc thực hiện những mục tiêu đó bằng niềm tin là rất hữu ích.

Sự trả giá hiếm khi rõ ràng, và quá trình để hiểu những sự trả giá có thể rất tốn thời gian cũng như tiền bạc. Tuy nhiên, sự phát triển của máy dự đoán đã giúp hiểu được logic và động lực cho những giá trị trả giá này.

Về mặt kinh tế, chi phí của việc xác định những sự trả giá chủ yếu sẽ là thời gian. Hãy xem xét một lộ trình cụ thể mà nhờ đó bạn có thể xác định những sự trả giá: cân nhắc và suy nghĩ. Hãy suy nghĩ điều bạn thực sự muốn đạt được hoặc chi phí cho sự không hài lòng của khách hàng có thể sẽ ra sao. Hãy dành thời gian để suy nghĩ, đối chiếu và có thể là hỏi lời khuyên của người khác. Hoặc có thể là dành thời gian nghiên cứu để hiểu rõ hơn về những sự trả giá.

Đối với việc phát hiện lừa đảo thẻ tín dụng, việc nghĩ về sự trả giá cho những khách hàng hài lòng hoặc không hài lòng và chi phí của việc cho phép một giao dịch lừa đảo được tiến hành là những bước đầu tiên rất cần thiết. Cung cấp những sự trả giá khác nhau cho những khách hàng có giá trị tài sản ròng lớn đòi hỏi nhiều sự suy nghĩ hơn. Việc đánh giá liệu những sự trả giá này có thay đổi khi những khách hàng đó trong kỳ nghỉ hay không, đòi hỏi sự cân nhắc kỹ lưỡng hơn. Vậy còn những khách hàng bình thường đang đi nghỉ thì sao? Sự trả giá trong những tình huống đó liệu có khác? Và có đáng để tách biệt việc du lịch trong chuyến công tác và kỳ nghỉ không? Hay những chuyến đi tới thành phố Rome từ Grand Canyon?

Trong mỗi trường hợp, việc đánh giá sự trả giá đòi hỏi thời gian và công sức: nhiều kết quả hơn đồng nghĩa với nhiều sự đánh giá hơn, đồng thời tốn nhiều thời gian và công sức hơn. Con người trải qua những chi phí nhận thức của sự đánh giá như là một quá trình đưa ra quyết định chậm. Chúng ta đều cần phải

quyết định chúng ta muốn trả giá bao nhiêu cho chi phí của sự trì hoãn quyết định. Một vài người sẽ chọn không nghiên cứu sâu những sự trả giá cho những bối cảnh có vẻ xa xôi hoặc ít có khả năng xảy ra. Mạng lưới thẻ tín dụng có thể sẽ thấy đáng để tách biệt việc du lịch trong chuyến đi công tác và kỳ nghỉ nhưng không phải là những kỳ nghỉ từ Grand Canyon đến thành phố Rome.

Trong những tình huống không khả thi như vậy, mạng lưới thẻ có thể sẽ đoán những quyết định chính xác, gộp vài thứ vào cùng với nhau, hoặc đơn giản là chọn một mặc định an toàn hơn. Nhưng với những quyết định thường xuyên hơn (ví dụ như việc đi du lịch nói chung) hoặc những điều có vẻ quan trọng hơn (ví dụ như những khách hàng có giá trị tài sản ròng lớn), mạng lưới thẻ sẽ mất thời gian để cân nhắc và xác định những sự trả giá cẩn thận hơn. Nhưng nó càng tốn nhiều thời gian để thử nghiệm thì quá trình đưa ra quyết định sẽ càng mất nhiều thời gian để trở nên hoàn hảo.

Xác định những sự trả giá cũng giống như nếm thử đồ ăn mới: thử một cái gì đó và xem điều gì xảy ra. Hoặc, nói theo thuật ngữ của kinh doanh hiện đại: thử nghiệm. Mỗi cá nhân có thể thực hiện những hành động khác nhau trong cùng một trường hợp và học được kết quả thực tế là gì. Họ học từ những sự trả giá thay vì nghĩ về chúng. Đương nhiên, bởi vì sự thử nghiệm đồng nghĩa với việc làm những gì mà bạn có thể sẽ coi là sai sót, những thử nghiệm cũng có chi phí. Bạn sẽ thử món ăn mà có thể bạn không thích. Nếu bạn tiếp tục thử những món ăn mới với hy vọng tìm được món gì đó lý tưởng, bạn đang bỏ lỡ nhiều bữa ăn ngon miệng. Sự đánh giá, cho dù là bằng sự cân nhắc hay sự thử nghiệm, đều tốn kém.

Biết lý do vì sao bạn làm những điều cụ thể

Sự dự đoán là cốt lõi của bước tiến đến với những chiếc xe tự lái, và sự phát triển của những nền tảng như Uber hay Lyft: lựa chọn tuyến đường giữa điểm đi và điểm đến. Những thiết bị điều hướng xe đã xuất hiện từ một vài thập kỷ, được xây dựng thành những chiếc xe hơi hoặc chỉ là những thiết bị độc lập. Nhưng sự phát triển của những thiết bị di động được kết nối Internet đã thay đổi dữ liệu mà phần mềm điều hướng của những nhà cung cấp nhận được. Ví dụ, trước khi sáp nhập vào Google, công ty khởi nghiệp Israel Waze đã tạo ra những bản đồ giao thông chính xác bằng việc theo dõi những tuyến đường mà người lái xe chọn. Sau đó nó sử dụng thông tin đó để cung cấp sự tối ưu

hóa hiệu quả về con đường nhanh nhất giữa hai điểm, dựa vào thông tin nó nhận được từ những người lái xe cũng như sự theo dõi liên tục giao thông. Nó cũng có thể dự đoán cách mà những điều kiện giao thông có thể tiếp diễn nếu bạn đi xa hơn, đồng thời có thể cung cấp những con đường mới và hiệu quả hơn trên tuyến đường nếu những điều kiện thay đổi. Những người dùng ứng dụng như Waze không phải lúc nào cũng làm theo chỉ dẫn. Họ không phản đối sự dự đoán, nhưng mục tiêu của họ có thể bao gồm nhiều yếu tố khác chứ không chỉ có mỗi tốc độ. Ví dụ, ứng dụng không biết người lái xe đang hết nhiên liệu và cần trạm xăng. Nhưng những người lái xe, biết rằng họ cần xăng, có thể bỏ qua sự gợi ý của ứng dụng và đi theo tuyến đường khác.

Tất nhiên là những ứng dụng như Waze có thể và sẽ trở nên tốt hơn. Ví dụ, như với xe Tesla chạy bằng điện thì sự điều hướng được dựa vào nhu cầu cần nạp điện và vị trí của những trạm điện. Một ứng dụng có thể chỉ đơn giản hỏi bạn liệu bạn có cần nhiên liệu hay không, hoặc trong tương lai, thậm chí có thể lấy dữ liệu trực tiếp từ xe của bạn. Điều này dường như là một vấn đề có thể giải quyết được, cũng giống như bạn có thể điều chỉnh cài đặt trong những ứng dụng điều hướng để tránh những đường có trạm thu phí.

Những yếu tố khác mà bạn ưa thích sẽ khó để lập trình hơn. Ví dụ, trên một đoạn đường dài, bạn có thể muốn đảm bảo rằng bạn đi qua những địa điểm phù hợp nhất định để nghỉ ngơi và ăn. Hoặc những tuyến đường nhanh nhưng chỉ tiết kiệm được 1-2 phút và thường rất tốn công khi lái. Hoặc bạn có thể không thích đi những con đường nhiều gió. Một lần nữa, những ứng dụng có thể sẽ học hỏi từ những hành vi, nhưng ở một thời điểm bất kỳ, một vài yếu tố sẽ không phải là một phần của sự dự đoán đã được mã hóa để tự động hóa một hành động. Máy móc có những hạn chế cơ bản về việc nó sẽ tốn bao nhiêu để có thể học cách dự đoán những sự ưu tiên của bạn.

Một điểm nhìn bao quát hơn cho những quyết định là những mục tiêu hiếm khi chỉ có một chiều. Con người sở hữu kiến thức, rõ ràng và hàm ẩn, về lý do họ làm một điều gì đó, điều đó có thể cho họ những cân nhắc cá nhân và chủ quan. Trong khi máy dự đoán điều gì có thể xảy ra, con người vẫn sẽ quyết định thực hiện hành động dựa trên sự hiểu biết của họ về mục tiêu. Trong nhiều tình huống, như với Waze, máy sẽ cho con người một sự dự đoán ngụ ý một kết quả nhất định một chiều (ví dụ như tốc độ); con người sau đó sẽ quyết định có nên bỏ qua sự gợi ý hành động đó hay không. Dựa

vào độ tinh xảo của máy dự đoán, con người có thể đòi hỏi nó cho một sự dự đoán khác dựa vào một sự ràng buộc mới ("Waze, hãy chỉ tôi qua chỗ một trạm xăng").

Sự dự đoán khó mã hóa

Ada Support, một công ty khởi nghiệp, sử dụng AI dự đoán để biến những câu hỏi hỗ trợ kỹ thuật khó thành dễ. AI trả lời những câu hỏi dễ và rồi gửi những câu hỏi khó tới con người. Với một nhà cung cấp dịch vụ, khi khách hàng gọi điện hỏi sự giúp đỡ, đa số những câu hỏi khách hàng hỏi đều đã được trả lời bởi những người khác. Hành động gõ câu trả lời thật dễ dàng. Nhưng thử thách ở đây là dự đoán khách hàng muốn gì và đánh giá xem nên cung cấp câu trả lời nào.

Thay vì hướng mọi người đến mục "những câu hỏi thường được hỏi" của một trang web, Ada xác định và trả lời những câu hỏi thường được hỏi ngay lập tức. Nó có thể sử dụng những đặc tính cá nhân của người dùng (ví dụ như kiến thức đã có về khả năng kỹ thuật, loại điện thoại họ đang dùng để gọi, hoặc những cuộc gọi trước đây) để cải thiện sự đánh giá của nó về câu hỏi. Trong quá trình, nó có thể giảm thiểu sự bực tức của người dùng, nhưng quan trọng hơn, nó có thể xử lý được nhiều sự tương tác một cách nhanh chóng mà không cần chi tiêu tiền vào những tổng đài sử dụng nhân lực tốn kém. Con người chuyên về giải quyết những câu hỏi không phổ biến và khó hơn, trong khi máy có thể xử lý những câu hỏi dễ.

Khi sự dự đoán của máy cải thiện, nó sẽ đánh giá nhiều hơn để xác định sự dự đoán trong nhiều tình huống. Giống như khi chúng ta giải thích suy nghĩ của mình cho người khác, chúng ta có thể giải thích suy nghĩ của chúng ta cho máy – dưới dạng mã phần mềm. Khi chúng ta mong chờ nhận được sự dự đoán chính xác, chúng ta có thể mã hóa sự dự đoán khó trước khi máy dự đoán. Ada làm được như vậy với những câu hỏi dễ. Sẽ tốn rất nhiều thời gian để đối phó với quá nhiều tình huống có thể xảy ra, vậy nên với những câu hỏi khó, Ada cần sự đánh giá của con người.

Kinh nghiệm đôi khi có thể giúp mã hóa khả năng dự đoán. Nhiều kinh nghiệm là vô hình và vì vậy không thể viết hay diễn giải dễ dàng được. Như Andrew McAfee và Erik Brynjolfsson từng viết: "Sự thay thế (con người bằng máy tính) bị ràng buộc vì có rất nhiều công việc mà con người ngầm

hiểu và hoàn thành một cách dễ dàng, nhưng ngay cả những lập trình viên hay bất kỳ ai cũng không thể giải thích rõ ‘quy tắc’ hay trình tự một cách rõ ràng.”[1] Tuy nhiên, điều đó không đúng với tất cả những công việc. Với một số quyết định, bạn có thể nêu rõ sự đánh giá bắt buộc và diễn giải nó bằng mã. Xét cho cùng, chúng ta thường hay giải thích suy nghĩ của mình cho người khác. Trên thực tế, mã hóa sự đánh giá cho phép bạn bổ sung vào phần “thì” trong mệnh đề ”nếu-thì”. Khi điều này xảy ra, sự đánh giá khi đó có thể được ghi nhận và được lập trình.

Thử thách ở đây là ngay cả khi bạn có thể lập trình sự đánh giá để thay thế con người, sự dự đoán mà máy dự đoán nhận được phải tương đối chính xác. Khi có quá nhiều tình huống có thể xảy ra, máy sẽ tốn rất nhiều thời gian để xác định cần làm gì trước trong từng tình huống. Bạn có thể dễ dàng lập trình máy để thực hiện một hành động cụ thể khi điều gì có vẻ đúng; tuy nhiên, khi vẫn còn sự không chắc chắn, việc nói với máy cần làm gì đòi hỏi một sự cân nhắc chi phí sai sót cẩn thận hơn. Sự không chắc chắn đồng nghĩa với việc bạn cần sự đánh giá khi dự đoán sai, chứ không chỉ khi dự đoán đúng. Nói cách khác, sự không chắc chắn sẽ làm tăng chi phí dự đoán sự trả giá cho một quyết định cụ thể.

Kỹ thuật chức năng phần thưởng (Reward Function Engineering – RFE)

Khi máy dự đoán cung cấp những dự đoán tốt hơn và có giá thành rẻ hơn, chúng ta cần tính toán để tận dụng những dự đoán này một cách tốt nhất. Cho dù là chúng ta có thể dự đoán trước hay không, ai đó sẽ cần xác định sự đánh giá. Đó là công việc của RFE , công việc xác định những phần thưởng của những hành động, khi sự dự đoán do AI đưa ra. Để làm tốt việc này đòi hỏi sự hiểu biết về những nhu cầu của tổ chức và những khả năng của máy.

Đôi khi kỹ thuật chức năng phần thưởng bao gồm cả sự đánh giá khó mã hóa – lập trình những phần thưởng trước khi dự đoán để tự động hóa những hành động. Những phương tiện tự lái là một ví dụ của những thành tựu khó mã hóa. Một khi sự dự đoán được thực hiện, hành động sẽ xảy ra ngay lập tức. Nhưng để đạt được phần thưởng ngay lập tức thì không hề dễ. RFE cần phải xem xét khả năng AI tối ưu hóa quá mức một số liệu thành công nào đó và khi làm như vậy nó sẽ hành động không nhất quán với những mục tiêu lớn hơn của tổ chức. Toàn bộ ủy ban đang nghiên cứu điều này cho xe lái tự

động; tuy nhiên, những sự phân tích như vậy sẽ đòi hỏi một loạt những quyết định mới.

Trong những trường hợp khác, dự đoán của những điều có thể xảy ra có thể khá tốn kém đối với bất kỳ ai muốn đánh giá trước toàn bộ những sự trả giá. Thay vào đó, con người cần đợi sự dự đoán và rồi đánh giá sự trả giá, mô hình đó gần giống với quá trình đưa ra quyết định, cho dù nó có bao gồm những sự dự đoán do máy tạo ra hay không. Như chúng ta sẽ thấy trong chương tiếp theo, máy móc đang dần xâm lấn vào lĩnh vực này. Trong một số trường hợp, máy dự đoán có thể học cách dự đoán sự đánh giá của con người thông qua việc quan sát những quyết định đã có.

Ghép chúng lại với nhau

Đa số chúng ta đều đã thử một vài kỹ thuật chức năng phần thưởng, nhưng là cho con người, chứ không phải với máy móc. Bố mẹ dạy con những giá trị. Những người cố vấn dạy những nhân viên mới cách hệ thống hoạt động. Những nhà quản lý đưa ra những mục tiêu cho nhân viên và rồi thúc đẩy họ làm việc hiệu suất cao hơn. Mỗi ngày, chúng ta đều đưa ra quyết định và đánh giá những thành tựu đạt được. Nhưng khi chúng ta làm như vậy với con người, chúng ta phân nhóm dự đoán và đánh giá cùng nhau, như vậy, vai trò của RFE là không nổi bật. Khi máy móc dự đoán tốt hơn, vai trò của RFE sẽ trở nên quan trọng.

Để minh họa RFE trên thực tế, hãy cùng xem xét những quyết định về chi phí ở ZipRecruiter, một trang công việc trực tuyến. Các công ty trả tiền cho ZipRecruiter để tìm những ứng viên đạt yêu cầu cho những cơ hội việc làm họ mong muốn được lấp đầy. Sản phẩm chính của ZipRecruiter là thuật toán phù hợp hoạt động hiệu quả và phát triển, một phiên bản của những công ty săn đầu người chuyên ghép cặp những người tìm việc làm với công ty.[2]

ZipRecruiter từng không chắc chắn về việc họ nên thu phí những công ty như thế nào cho dịch vụ của mình. Thu phí ít thì sẽ không có được nhiều tiền. Thu phí nhiều thì khách hàng sẽ chuyển qua công ty của các đối thủ. Để xác định chi phí, ZipRecruiter đã tìm gặp hai chuyên gia, J.P. Dubé và Sanjog Misra, những chuyên gia kinh tế học đến từ Trường Kinh doanh Booth của Đại học Chicago, những người đã thiết kế những thí nghiệm để xác định chi phí tốt nhất. Họ phân chia giá thành khác nhau cho những nhóm khách hàng khác

nhau và xác định khả năng mà mỗi nhóm có thể mua. Điều này cho phép họ quyết định xem những khách hàng khác nhau phản ứng như thế nào với những giá thành khác nhau.

Thử thách này là để tìm ra thứ "tốt nhất" là gì. Công ty có nên tối đa hóa doanh thu ngắn hạn? Để làm được điều này, công ty sẽ phải chọn đưa ra giá thành cao. Nhưng giá thành cao đồng nghĩa với việc ít khách hàng hơn (ngay cả khi mỗi khách hàng đều có thể đem lại lợi nhuận hơn). Điều đó đồng nghĩa với việc ít sự truyền miệng hơn. Bên cạnh đó, nếu họ có ít bài đăng công việc hơn, số người sử dụng ZipRecruiter để tìm việc có thể sẽ giảm. Cuối cùng, những khách hàng đối mặt với giá thành cao có thể sẽ bắt đầu tìm kiếm những lựa chọn khác. Trong khi họ có thể phải trả giá cao trong ngắn hạn, họ có thể chuyển sang một công ty đối thủ khác trong thời gian dài. Vậy ZipRecruiter nên cân nhắc những yếu tố đa dạng này như thế nào? Họ nên tối đa hóa sự trả giá nào?

Khá đơn giản để có thể tính toán những hệ quả trong thời gian ngắn hạn của việc tăng giá thành. Các chuyên gia cho rằng sự tăng giá thành với một số loại khách hàng mới có thể sẽ tăng lợi nhuận ngày qua ngày đến khoảng 50%. Tuy nhiên, ZipRecruiter không hành động ngay lập tức. Họ nhận ra rủi ro trong dài hạn và chờ xem những khách hàng trả giá cao hơn có rời đi hay không. Sau bốn tháng, họ nhận ra rằng việc tăng giá thành vẫn sinh lợi nhuận cao. Họ không muốn bỏ qua việc có lợi nhuận cao nữa và đánh giá bốn tháng là đủ dài để thực hiện những sự thay đổi giá thành.

Xác định những thành tựu từ những hành động đa dạng – mảnh ghép quan trọng của sự đánh giá – chính là tính năng của kỹ thuật chức năng phần thưởng một phần quan trọng trong quá trình đưa ra quyết định của con người. Máy dự đoán là công cụ cho con người. Khi con người cần xác định những kết quả và áp đặt sự đánh giá, máy dự đoán đóng vai trò quan trọng khi nó cải thiện hơn.

NHỮNG ĐIỂM CHÍNH

• Máy dự đoán làm tăng tính quan trọng của sự đánh giá bởi vì khi sự dự đoán trở nên rẻ hơn, việc hiểu những thành tựu liên quan đến những hành động sẽ tăng giá trị. Tuy nhiên, sự đánh giá là rất tốn kém. Xác định được những sự trả giá tương quan cho những hành động khác nhau trong những

tình huống khác nhau tốn thời gian, nỗ lực và sự thử nghiệm.

• Nhiều quyết định xảy ra dưới những điều kiện không chắc chắn. Chúng ta quyết định mang theo ô vì chúng ta nghĩ trời có thể mưa, nhưng chúng ta có thể sai. Chúng ta quyết định đồng ý một giao dịch bởi vì chúng ta nghĩ nó hợp pháp, nhưng chúng ta có thể sai. Dưới những điều kiện của sự không chắc chắn, chúng ta cần phải xác định sự trả giá cho việc hành động dựa theo những quyết định sai lầm, chứ không phải là những quyết định đúng. Vậy nên sự không chắc chắn làm gia tăng chi phí của việc đánh giá sự trả giá của một quyết định cụ thể.

• Nếu có thể quản lý một số tình huống hành động kết hợp với một quyết định, thì chúng ta có thể chuyển sự đánh giá cho máy dự đoán (đây là RFE) để máy có thể tự đưa ra quyết định một khi nó tạo ra sự dự đoán. Điều này cho phép tự động hóa quyết định. Tuy nhiên, thường có quá nhiều tình huống hành động, nên nó quá tốn kém để mã hóa tất cả những sự trả giá liên quan đến mỗi sự kết hợp, đặc biệt là những sự kết hợp hiếm hoi. Trong những trường hợp này, sẽ hiệu quả hơn nếu con người áp dụng sự đánh giá sau khi máy dự đoán hoạt động.

9Dự đoán sự đánh giá

N

hững công ty như chi nhánh của Google – Waymo đã thí nghiệm thành công cách tự động hóa việc di chuyển con người giữa hai điểm. Nhưng đó chỉ là một phần trong việc tạo ra những phương tiện tự động. Việc lái xe cũng có ảnh hưởng đến những hàng khách trong xe, điều này khó để quan sát hơn. Những người lái xe, tuy nhiên, có quan tâm đến những người khách trong xe. Một trong những điều đầu tiên mà người mới học lái xe học là phanh xe sao cho những người trong xe thấy thoải mái. Những chiếc xe của Waymo phải được dạy để tránh dừng đột ngột và thay vào đó là phanh từ từ.

Có hàng nghìn những quyết định liên quan trong việc lái xe.[1] Để con người tự mã hóa sự đánh giá của mình về cách xử lý mỗi tình huống có thể xảy ra là không thực tế. Thay vào đó, chúng ta đào tạo những hệ thống lái xe tự động bằng cách cho chúng thấy nhiều ví dụ để chúng học cách dự đoán sự đánh giá của con người: "Con người sẽ làm gì trong tình huống này?" Lái xe không phải chuyện gì đặc biệt. Trong bất kỳ môi trường nào mà con người đưa ra quyết định liên tục, chúng ta có thể thu thập dữ liệu về những gì họ nhận được và những quyết định họ đưa ra, chúng ta sẽ có thể tự động hóa những quyết định đó bằng việc thưởng cho máy dự đoán để dự đoán: *Con người sẽ làm gì?*

Một câu hỏi quan trọng, ít nhất là với con người, rằng liệu AI có thể thay đổi sức mạnh dự đoán của nó và dần dần không còn cần tới con người nữa hay không?

Lý giải con người

Nhiều quyết định phức tạp và được dự đoán dựa trên sự đánh giá khó mã hóa. Tuy nhiên, điều này không đảm bảo rằng con người vẫn sẽ là một phần quan trọng của những quyết định này. Thay vào đó, như với những chiếc xe tự động, máy có thể học cách dự đoán sự đánh giá của con người thông qua việc quan sát những ví dụ. Vấn đề dự đoán trở thành: "Nếu được cho dữ liệu đầu vào, con người sẽ làm gì?"

Công ty Grammarly là một ví dụ. Thành lập vào năm 2009 bởi Alex Shevchenko và Max Lytvyn, Grammarly tiên phong trong việc sử dụng máy tự học để cải thiện cấu trúc những tài liệu viết có văn phong trang trọng. Sự tập trung chính của nó là vào việc cải thiện ngữ phép và chính tả trong câu. Ví dụ, đặt câu vừa rồi vào Grammarly, và nó sẽ nói cho bạn biết rằng "Sự tập trung của nó" nên sửa thành "Nó tập trung" và "ngữ phép" đã bị đánh vần sai (nó nên là "ngữ pháp" mới đúng). Nó cũng sẽ nói cho bạn biết rằng từ "chính" thường bị sử dụng quá đà.

Grammarly đã có được những sự chỉnh sửa này bằng việc nghiên cứu một kho tài liệu mà những biên tập viên có tay nghề cao đã chỉnh sửa và thông qua việc học hỏi từ những phản hồi của người dùng chấp nhận hay từ chối những đề xuất. Trong cả hai trường hợp, Grammarly đã dự đoán việc một người biên tập sẽ làm. Nó vượt xa những ứng dụng cơ học của những quy tắc ngữ pháp để đánh giá được rằng những sự sai lệch ngữ pháp nào được ưa thích hơn bởi người đọc.

Ý tưởng rằng con người có thể đào tạo AI mở rộng trong nhiều tình huống. AI trung tâm của Lola, một công ty khởi nghiệp tự động hóa quá trình đặt phòng du lịch, bắt đầu bằng việc tìm kiếm những lựa chọn khách sạn tốt. Nhưng, như *New York Times* đưa tin:

Nó không thể cùng đẳng cấp về mặt chuyên môn, ví dụ, của một đại diện với nhiều năm kinh nghiệm đặt phòng nghỉ cho gia đình tới Disney World. Con người có thể nhanh nhẹn hơn, ví dụ, họ biết cách tư vấn cho một gia đình với niềm hy vọng có được một bức ảnh không bị che khuất với những đứa trẻ đứng trước Lâu đài Lọ Lem rằng họ nên đặt trước một bàn ăn sáng ở trong công viên, trước giờ mở cổng.[2]

Ví dụ này cho thấy máy có thể dễ dàng áp dụng sự đánh giá ở những trường hợp được miêu tả rõ (ví dụ, phòng trống và giá thành), nhưng không thể hiểu được những ưu tiên tinh tế hơn của con người. Tuy nhiên, Lola có thể học cách dự đoán việc mà con người với mức độ kinh nghiệm cao và suy nghĩ có thể làm. Câu hỏi đặt ra cho Lola là: Cần quan sát bao nhiêu người đặt phòng nghỉ tại Orlando để máy dự đoán có đủ phản hồi và học từ những tiêu chí liên quan khác? Như Lola nhận ra rằng, mặc dù AI của công ty bị thách thức bởi một số tiêu chí, nó có thể phát hiện ra những quyết định mà người đại diện đã

thực hiện nhưng không thể miêu tả trước được, ví dụ như việc chọn những khách sạn hiện đại hoặc những khách sạn ở góc phố theo sở thích. Người huấn luyện giúp AI trở nên tốt đến mức con người dần trở nên không cần thiết trong nhiều khía cạnh của công việc. Điều này đặc biệt quan trọng khi AI tự động hóa một quy trình không chấp nhận sự sai sót. Con người có thể giám sát AI và sửa chữa sai sót. Theo thời gian, AI học từ sai sót của nó cho đến khi sự chỉnh sửa của con người không còn cần thiết.

Một ví dụ khác là X.ai, một công ty khởi nghiệp chuyên cung cấp một người trợ lý có thể sắp xếp các lịch hẹn và đưa chúng vào lịch trình của bạn.[3] Nó tương tác với người dùng và những người mà người dùng đó muốn gặp bằng email thông qua một trợ lý cá nhân kỹ thuật số (“Amy” hoặc “Andrew”, tùy vào sở thích của bạn). Ví dụ, bạn có thể gửi email đến Andrew để sắp xếp một lịch hẹn giữa bạn và ông H vào thứ 5 tuần sau. X.ai sẽ đánh giá lịch trình của bạn và gửi email đến ông H để đặt lịch hẹn. Ông H có thể không ngờ rằng Andrew không phải là con người. Vấn đề là bạn đã được giải phóng khỏi việc giao tiếp với ông H hoặc trợ lý của ông ta (có thể là một Amy hoặc Andrew khác).

Rõ ràng, thảm họa có thể xảy ra nếu sai sót trong việc lập kế hoạch xảy ra hoặc nếu trợ lý tự động xúc phạm đến người mời tiềm năng. Trong nhiều năm, X.ai tuyển dụng những người huấn luyện. Họ đánh giá độ chính xác từ những phản hồi của AI và xác thực chúng. Mỗi khi người huấn luyện thực hiện một sự thay đổi, AI học hỏi để đưa ra một phản hồi tốt hơn.[4] Vai trò của người huấn luyện không chỉ là đảm bảo sự lịch sự đúng mực, họ cũng phải xử lý những người có hành vi xấu với trợ lý ảo.[5] Tính đến thời điểm viết, câu hỏi rằng cách tiếp cận đánh giá sự dự đoán này có thể tự động hóa đến mức nào vẫn chưa được giải đáp.

Liệu con người có bị cho ra rìa?

Nếu máy có thể học để dự đoán hành vi của con người, liệu con người có bị cho ra rìa hoàn toàn? Với quỹ đạo hiện tại của máy dự đoán, chúng tôi không nghĩ vậy. Con người là một nguồn tài nguyên, nên những nguyên lý kinh tế đơn giản gợi ý rằng họ sẽ vẫn làm gì đó. Câu hỏi đặt ra là “cái gì đó” của con người có giá trị cao hay thấp, hấp dẫn hay không hấp dẫn. Con người trong tổ chức của bạn nên làm gì? Bạn nên tìm kiếm những gì khi tuyển dụng mới?

Sự dự đoán phụ thuộc vào dữ liệu. Điều đó đồng nghĩa với việc con người có hai lợi thế so với máy. Chúng ta biết một vài điều mà máy không biết (chưa biết) và quan trọng là chúng ta giỏi trong việc quyết định cần làm gì khi không có nhiều dữ liệu.

Con người có ba loại dữ liệu mà máy không có. Đầu tiên, các giác quan của con người rất mạnh mẽ. Theo nhiều cách, mắt, tai, mũi và da vẫn vượt qua những năng lực của máy. Thứ hai, con người là người quyết định tối cao với những sở thích của chính họ. Dữ liệu của người tiêu dùng cực kỳ có giá trị bởi nó cho máy dự đoán dữ liệu về những sở thích này. Những cửa hàng tạp hóa cung cấp các phiếu giảm giá cho người tiêu dùng sử dụng thẻ khách hàng thân thiết để thu thập dữ liệu về hành vi của họ. Các cửa hàng trả tiền người tiêu dùng để họ tiết lộ những sở thích của họ. Google, Facebook và những trang khác cung cấp những dịch vụ miễn phí để đổi lấy dữ liệu mà họ có thể sử dụng trong nhiều bối cảnh khác nhau để đặt mục tiêu quảng cáo. Thứ ba, các vấn đề liên quan đến quyền riêng tư hạn chế dữ liệu có sẵn cho máy. Nếu nhiều người lựa chọn không tiết lộ hoạt động tình dục, tình hình tài chính, tình trạng sức khỏe tâm lý và những suy nghĩ không phù hợp, máy dự đoán sẽ không có đủ dữ liệu để dự đoán nhiều loại hành vi. Trong trường hợp thiếu dữ liệu tốt, sự hiểu biết của chúng ta về những người khác sẽ đóng vai trò quan trọng đối với sự đánh giá.

Dự đoán với ít dữ liệu

Máy dự đoán có thể cũng thiếu dữ liệu bởi vì một số sự kiện rất hiếm xảy ra. Nếu máy không quan sát đủ những quyết định của con người, nó không thể dự đoán sự đánh giá ẩn sau những quyết định này.

Trong chương 6, chúng ta đã thảo luận về “những điều chưa biết là chưa biết”, những sự kiện hiếm khi xảy ra khó để dự đoán vì thiếu dữ liệu, bao gồm những cuộc bầu cử tổng thống và những trận động đất. Ở một số trường hợp, con người giỏi việc dự đoán với ít dữ liệu, ví dụ chúng ta có thể nhận ra những khuôn mặt ngay cả khi người đó già đi. Chúng ta cũng đã thảo luận rằng “những điều chưa biết là chưa biết”, theo định nghĩa là rất khó để dự đoán hoặc phản ứng. AI không thể dự đoán con người sẽ làm gì nếu con người chưa từng đối mặt với một tình huống tương tự. Bằng cách này, AI không thể dự đoán định hướng chiến lược của một công ty trước một công nghệ mới, ví dụ như Internet, công nghệ sinh học và kể cả bản thân AI. Con

người có thể đưa ra những sự so sánh hoặc nhận ra những sự giống nhau có ích trong những hoàn cảnh khác nhau.

Cuối cùng, máy dự đoán có thể sẽ giỏi việc so sánh hơn. Tuy vậy, quan điểm của chúng tôi rằng máy dự đoán sẽ dự đoán không tốt những sự kiện hiếm khi xảy ra vẫn còn đó. Trong tương lai có thể nhìn thấy được, con người sẽ có vai trò trong việc dự đoán và đánh giá khi những tình huống không phổ biến phát sinh.

Trong chương 6, chúng tôi cũng nhấn mạnh “những điều chưa biết là đã biết”. Ví dụ, chúng ta đã thảo luận những thách thức của việc quyết định có nên giới thiệu cuốn sách này cho bạn hay không, ngay cả khi bạn trở nên cực kỳ thành công trong việc quản lý AI trong tương lai. Thách thức ở đây là bạn không có dữ liệu về việc gì có thể xảy ra nếu bạn không đọc cuốn sách này. Nếu bạn muốn hiểu điều gì dẫn đến điều gì, bạn cần phải quan sát điều gì sẽ xảy ra trong tình huống đối lập.

Con người có thể cung cấp hai giải pháp chính cho vấn đề này: thí nghiệm và mô hình hóa. Nếu tình huống phát sinh đủ thường xuyên, bạn có thể chạy thử một thử nghiệm kiểm soát ngẫu nhiên. Chỉ định một số người thử nghiệm (bắt họ đọc cuốn sách, hoặc ít nhất là đưa họ cuốn sách và có thể tổ chức một vài kì thi liên quan) và những người còn lại kiểm soát (bắt họ không đọc cuốn sách, hoặc đừng quảng cáo cuốn sách cho họ). Hãy chờ đợi và thu thập những biện pháp họ áp dụng AI vào công việc của họ. So sánh hai nhóm. Sự khác biệt giữa nhóm thử nghiệm và nhóm kiểm soát chính là hiệu quả của việc đọc sách. Những thí nghiệm như vậy có ảnh hưởng mạnh mẽ. Không có chúng, những phương pháp điều trị y khoa mới sẽ không được chấp nhận. Chúng tiếp năng lượng cho nhiều quyết định ở những công ty được thúc đẩy bởi dữ liệu từ Google đến Capital One.

Mô hình hóa, một sự thay thế cho việc thí nghiệm, bao gồm sự hiểu biết rõ về tình huống và quá trình đưa ra dữ liệu để quan sát. Nó đặc biệt hữu ích khi những thí nghiệm là bất khả thi bởi tình huống không phát sinh đủ thường xuyên hoặc chi phí cho một thí nghiệm quá cao.

Quyết định của trang công việc trực tuyến ZipRecruiter để tìm giá thành tốt nhất, mà chúng tôi đã mô tả trong chương trước, bao gồm hai phần. Đầu tiên, họ cần phải xác định “điều tốt nhất” là gì: doanh thu ngắn hạn hay dài hạn?

Thứ hai, họ cần chọn một giá thành cụ thể. Để giải quyết vấn đề thứ hai, họ đã tiến hành thử nghiệm. Những chuyên gia đã thiết kế thử nghiệm, nhưng về nguyên tắc, khi AI cải thiện, với đủ quảng cáo và thời gian, những thử nghiệm như vậy có thể được tự động hóa. Tuy nhiên, việc xác định "điều tốt nhất" khó tự động hóa hơn. Bởi vì số lượng người tìm việc làm phụ thuộc vào số lượng quảng cáo công việc và ngược lại, thị trường tổng thể chỉ có một quan sát. Nếu làm sai, ZipRecruiter có thể không kinh doanh được và sẽ không có cơ hội thứ hai. Vậy nên, họ đã mô hình hóa doanh nghiệp của mình. Họ khám phá những hệ quả của việc tối đa hóa lợi nhuận ngắn hạn và so sánh nó với những mô hình tương tự với mục tiêu là tối đa hóa lợi nhuận dài hạn. Không có dữ liệu, những kết quả mô hình và kỹ thuật chức năng phần thưởng vẫn sẽ được thực hiện của con người - những người có tay nghề cao.

Mô hình hóa cũng giúp phe Đồng minh ném bom trong Thế chiến II. Các kỹ sư nhận ra rằng họ có thể trang bị những máy bay ném bom tốt hơn. Cụ thể, họ có thể thêm khối lượng vào máy bay mà không làm ảnh hưởng đến máy bay. Câu hỏi đặt ra là điểm cụ thể nào cần bảo vệ. Họ có thể thử nghiệm, nhưng như vậy rất tốn kém. Những phi công có thể sẽ mất mạng. Với mỗi chiếc máy bay ném bom trở về sau những trận ném bom trên khắp nước Đức, các kỹ sư có thể nhìn thấy nơi chúng bị bắn bởi đạn diệt máy bay. Các lỗ đạn trên thân máy bay là dữ liệu. Nhưng đây liệu có phải là những nơi cần để bảo vệ?

Họ hỏi nhà thống kê Abraham Wald để đánh giá vấn đề này. Sau khi suy nghĩ và tính toán cẩn thận, ông nói với họ cần bảo vệ máy bay ở những nơi không có lỗ đạn. Liệu ông có nhầm không? Ý ông không phải là cần bảo vệ những khu vực có lỗ đạn sao? Không. Ông có một mô hình quá trình tạo ra dữ liệu. Ông nhận ra rằng một vài máy bay ném bom không trở về từ trận ném bom và phỏng đoán rằng chúng bị bắn trúng ở những chỗ hiểm. Ngược lại, những máy bay trở về bị bắn trúng ở những chỗ không gây nguy hiểm. Với cái nhìn này, các kỹ sư không quân đã tăng giáp ở những nơi không có lỗ đạn, và những máy bay này đã được bảo vệ tốt hơn.[6] Cái nhìn sâu sắc của Wald về những dữ liệu còn thiếu đòi hỏi sự hiểu biết về nơi bắt nguồn của dữ liệu; do vấn đề này chưa từng phát sinh, các kỹ sư không có những tiền lệ để đưa ra kết luận. Trong tương lai có thể nhìn thấy, những sự tính toán như vậy vượt ngoài khả năng của máy dự đoán.

Vấn đề này khó để giải quyết. Giải pháp đến từ con người, chứ không phải từ máy dự đoán. Tuy nhiên, con người là một trong những nhà thống kê giỏi nhất trong lịch sử. Con người có sự hiểu biết về khía cạnh toán học của số liệu thống kê và có tâm thế đủ linh hoạt để hiểu quá trình tạo ra dữ liệu. Con người có thể học những kỹ năng mô hình hóa như vậy qua đào tạo. Đây là khía cạnh chính của đa số những chương trình tiến sĩ kinh tế và là một phần của chương trình MBA tại nhiều trường (bao gồm những khóa học chúng tôi phát triển ở Đại học Toronto). Những kĩ năng như vậy quan trọng khi làm việc với máy dự đoán. Nếu không, rất dễ để rơi vào cái bẫy của những điều chưa biết là đã biết.

Giống như Wald có mô hình tốt về quá trình tạo ra dữ liệu về các lỗ đạn, một mô hình tốt về hành vi con người có thể giúp đưa ra những sự dự đoán chính xác hơn khi chúng tạo ra dữ liệu. Trong tương lai gần, con người cần giúp phát triển những mô hình như vậy và xác định những yếu tố dự báo liên quan đến hành vi. Máy dự đoán sẽ gặp khó khăn khi ngoại suy trong tình huống như vậy khi nó không có dữ liệu vì hành vi có khả năng sẽ thay đổi. Nó cần hiểu con người.[7]

Những vấn đề tương tự phát sinh trong nhiều quyết định liên quan đến câu hỏi, “Điều gì sẽ xảy ra nếu tôi làm điều này?”. Bạn có nên thêm một sản phẩm mới vào dòng sản phẩm không? Bạn có nên hợp nhất với đối thủ cạnh tranh? Bạn có nên mua một công ty khởi nghiệp sáng tạo hay một kênh đối tác không?[8] Nếu con người hành xử khác đi khi có sự thay đổi thì hành vi trong quá khứ không còn là chỉ dẫn hữu ích cho hành vi trong tương lai nữa. Máy dự đoán sẽ không có dữ liệu liên quan. Với những sự kiện hiếm khi xảy ra, máy dự đoán sẽ bị hạn chế. Những sự kiện hiếm khi xảy ra cung cấp một hạn chế quan trọng tới khả năng máy dự đoán những đánh giá của con người.

NHỮNG ĐIỂM CHÍNH

• Máy có thể học để dự đoán sự đánh giá của con người. Một ví dụ là lái xe. Việc mã hóa đánh giá của con người về cách xử lý trong mỗi tình huống có thể xảy ra là không thực tế. Tuy nhiên, chúng ta đào tạo những hệ thống lái xe tự động bằng cách cho chúng nhiều ví dụ và thưởng chúng vì đã dự đoán sự đánh giá của con người: Con người sẽ làm gì trong tình huống đó?

• Khả năng dự đoán sự đánh giá của con người của máy cũng có những hạn chế liên quan đến việc thiếu dữ liệu. Có một số loại dữ liệu mà con người có nhưng máy lại không, ví dụ những sở thích của con người. Những dữ liệu như vậy có giá trị và các công ty hiện đang trả tiền để được tiếp cận chúng thông qua các chương trình giảm giá, thẻ khách hàng thân thiết và những dịch vụ trực tuyến miễn phí như Google và Facebook.

• Máy không giỏi việc dự đoán những sự kiện hiếm khi xảy ra. Những người quản lý đưa ra những quyết định về việc sáp nhập, đổi mới và cộng tác mà không cần dữ liệu của những sự kiện tương tự trong quá khứ. Con người sử dụng sự so sánh và những mô hình để đưa ra quyết định trong những tình huống không phổ biến như vậy. Máy không thể dự đoán sự đánh giá khi một tình huống chưa từng xảy ra nhiều lần trong quá khứ.

10Chế ngự sự phức tạp

B

ộ phim truyền hình *The Americans* (tạm dịch: Cuộc chiến thầm lặng), lấy bối cảnh Washington DC trong thời kì Chiến tranh lạnh, có sự xuất hiện của một robot đưa thư và phân loại tài liệu khắp văn phòng FBI. Việc một cỗ máy tự động như vậy tồn tại vào những năm 1980 có vẻ đáng ngạc nhiên. Được tiếp thị như là Mailmobile, nó xuất hiện lần đầu tiên vào một thập kỷ trước.[1]

Để hướng dẫn Mailmobile, một kỹ thuật viên sẽ vẽ một con đường bằng chất hóa học phát ra tia cực tím ở thảm từ phòng thư đến nhiều văn phòng. Con robot sử dụng máy cảm biến để từ từ đi theo con đường đó (tối thiểu với tốc độ hơn 100 dặm một giờ) cho tới khi các dấu hiệu hóa học báo hiệu nó dừng lại. Mailmobile có trị giá từ khoảng 10.000 đến 12.000 đô la (khoảng 50.000 đô la theo thời giá hiện nay), và với một khoản phụ phí, công ty có thể gắn một máy cảm biến để phát hiện những chướng ngại vật trên đường của nó. Nếu không nó sẽ kêu bíp rất nhiều để báo với mọi người rằng nó đang đến. Trong một văn phòng mà con người mất đến hai tiếng để chuyển thư, Mailmobile hoàn thành công việc chỉ trong 20 phút vì không dừng lại để buôn chuyện văn phòng. Robot đưa thư đòi hỏi sự lên kế hoạch cẩn thận. Nó chỉ có thể xử lý một số lượng nhỏ thay đổi trong môi trường.

Đến tận ngày nay, nhiều hệ thống đường sắt tự động trên thế giới đã có sự yêu cầu lắp đặt bao quát. Ví dụ, tàu điện ngầm ở Copenhagen không cần người lái, nhưng nó hoạt động bởi các chuyến tàu hoạt động trong một môi trường đã được lên kế hoạch cẩn thận; chỉ có một vài máy cảm biến thông báo cho robot về môi trường của nó.

Những hạn chế này là một yếu tố phổ biến của đa số các máy và thiết bị. Chúng được thiết kế để hoạt động trong những môi trường bị giới hạn. So với đa số thiết bị ở nhà máy, robot đưa thư rất ấn tượng bởi vì nhiều văn phòng có thể cài đặt nó tương đối dễ dàng. Nhưng, phần lớn robot cần một môi trường được kiểm soát chặt chẽ và chuẩn hóa để hoạt động bởi thiết bị không chịu đựng được sự không chắc chắn.

Nhiều "nếu"

Đa số các máy – kể cả cứng hay mềm – đều được lập trình sử dụng logic kinh điển nếu-thì. Phần "nếu" xác định một viễn cảnh, một điều kiện môi trường, hoặc một phần thông tin. Phần "thì" nói cho máy biết cần làm gì cho mỗi "nếu" (và "nếu không" và "khác"): "Nếu con đường hóa học không còn được tìm thấy, thì dừng lại". Robot đưa thư không có khả năng nhìn thấy môi trường xung quanh nó và chỉ có thể hoạt động trong môi trường giảm sự "nếu" mà nó có thể đối phó một cách nhân tạo. Nếu nó có thể phân biệt giữa các tình huống nhiều "nếu" hơn và nếu nó không thay đổi điều nó làm, về cơ bản là dừng và đi ở bất kỳ điểm nào, nó có thể đã được sử dụng ở nhiều nơi hơn. Roomba của thời hiện đại – robot hút bụi tự động của iRobot – có thể làm được điều này và tự do lang thang khắp các phòng với máy cảm biến để ngăn chặn việc nó ngã xuống cầu thang hoặc bị mắc kẹt ở các góc, cùng với bộ nhớ để đảm bảo nó hút bụi sạch sàn một cách kịp thời.

Sự dự đoán tốt hơn xác định được nhiều "nếu" hơn. Với nhiều "nếu", robot đưa thư có thể phản ứng với nhiều tình huống hơn. Máy dự đoán cho phép robot đưa thư xác định môi trường trời tối và ướt cùng với con người đứng cách đó 20 feet và một chú mèo ở đằng trước đòi hỏi robot cần đi chậm lại, nhưng môi trường trời tối và ướt cùng với con người đứng cách đó 20 feet và một chú sóc ở đằng trước thì có thể không khiến robot đi chậm lại. Máy dự đoán cho phép robot di chuyển mà không cần đường ray đã được lên kế hoạch trước. Mailmobile mới có thể hoạt động trong nhiều môi trường mà không cần tốn thêm chi phí.

Robot vận chuyển xuất hiện rất nhiều. Các kho hàng có nhiều hệ thống đưa hàng tự động có thể dự đoán môi trường xung quanh và điều chỉnh sao cho phù hợp. Hạm đội robot của Kiva vận chuyển các sản phẩm bên trong những nhà máy của Amazon. Các công ty khởi nghiệp đang thử nghiệm với robot vận chuyển có thể đưa gói hàng (hoặc pizza) đi trên vỉa hè và đường phố từ các doanh nghiệp đến nhà khách hàng và quay về.

Robot ngày nay có thể làm được điều này vì chúng có thể sử dụng dữ liệu từ các máy cảm biến tinh vi để dự đoán môi trường và sau đó nhận hướng dẫn cách xử lý. Chúng ta không thường xuyên khái niệm hóa điều này thành dự đoán, nhưng về cơ bản thì là như vậy. Và nếu nó tiếp tục có giá thành rẻ hơn, nhiều robot sẽ trở nên tốt hơn.

Nhiều “thì”

George Stigler, một chuyên gia kinh tế nhận giải thưởng Nobel, đã nhận xét: “Ai mà chưa từng bị lỡ chuyến bay thì đã tốn rất nhiều thời gian ở sân bay.”[2]Khi một logic đặc biệt được đưa ra, một tranh luận phản bác cũng rất mạnh mẽ: bạn có thể hoàn thành công việc hoặc thư giãn một cách dễ dàng ở sân bay như ở những nơi khác, và bạn có thể yên tâm khi đến đó sớm để tránh lỡ chuyến bay. Do vậy, phòng chờ máy bay mới xuất hiện. Các hãng hàng không phát minh ra nó để cung cấp cho hành khách (hoặc ít nhất là những người giàu có hoặc thường xuyên bay) một không gian tiện lợi và yên tĩnh để chờ đợi chuyến bay của họ. Ai đến muộn sẽ chỉ sử dụng phòng chờ máy bay khi quá cảnh hoặc khi chuyến bay bị trì hoãn hoặc để khóc nếu họ bị lỡ chuyến bay tới Bali. Phòng chờ máy bay cung cấp không gian để thư giãn khi bạn đến sân bay không đúng giờ (điều này có khả năng xảy ra thường xuyên).

Giả sử bạn có một chuyến bay lúc 10 giờ sáng, hướng dẫn của hãng hàng không sẽ nói rằng bạn nên đến sớm 60 phút lúc 9 giờ sáng và vẫn kịp chuyến bay. Vậy thì mấy giờ bạn nên bắt đầu rời đi để đến sân bay? Bạn thường có thể đến sân bay trong vòng 30 phút, điều này cho phép bạn rời khỏi nhà lúc 8 giờ 30 phút sáng, nhưng giả định này chưa tính toán tắc nghẽn giao thông. Khi bay trở lại Toronto từ một cuộc hẹn ở New York về cuốn sách này, tình trạng giao thông tồi tệ ở sân bay LaGuardia đã khiến chúng tôi phải đi bộ dặm cuối cùng dọc đường cao tốc. Việc đó có thể dễ dàng làm mất thêm 30 phút (hoặc hơn, nếu bạn không thích rủi ro). Vậy là bạn chỉ còn sự lựa chọn là 8 giờ sáng, thời gian bạn nên bắt đầu rời đi nếu bạn không biết giao thông sẽ như thế nào. Kết quả là bạn thường xuyên phải chờ khoảng 30 phút hoặc hơn ở phòng chờ sân bay.

Những ứng dụng như Waze cung cấp thời gian di chuyển chính xác từ vị trí hiện tại của bạn đến sân bay. Những ứng dụng đó theo dõi cả những mẫu giao thông trong thời gian thực và trong quá khứ để vừa dự đoán, vừa cập nhật những tuyến đường nhanh nhất. Kết hợp nó với Google Now, bạn có thể dự đoán bất kỳ sự chậm trễ nào có thể xuất hiện trong chuyến bay của bạn với những ứng dụng theo dõi những sự trì hoãn chuyến bay trong lịch sử hoặc vị trí nơi máy bay kết nối khác.

Sự dự đoán tốt hơn, bằng việc giảm hoặc loại bỏ nguồn chính của sự không chắc chắn, loại bỏ nhu cầu cần nơi để chờ ở sân bay của bạn. Quan trọng hơn, sự dự đoán tốt hơn cho phép bạn có những hành động mới. Thay vì có quy tắc cứng nhắc là đi trước 2 tiếng, bạn có thể có một quy tắc ngẫu nhiên, thu thập những thông tin và nói cho bạn biết khi nào cần rời đi. Những quy tắc ngẫu nhiên này là những mệnh đề "nếu-thì" và cho phép nhiều sự "thì" hơn (rời đi sớm, đúng giờ, hoặc trễ), phụ thuộc vào những sự dự đoán ngày càng đáng tin. Vậy ngoài việc tạo ra nhiều "nếu", sự dự đoán tăng những cơ hội bằng việc tăng số lượng những "thì" có thể xảy ra.

Nhiều "nếu" và "thì" hơn

Sự dự đoán tốt hơn cho phép bạn dự đoán nhiều thứ thường xuyên hơn và giảm sự không chắc chắn. Mỗi dự đoán mới cũng có một ảnh hưởng gián tiếp: nó đưa ra những sự lựa chọn khả thi mà bạn chưa từng cân nhắc trước đây. Và bạn không cần phải mã hóa rõ ràng sự "nếu" và "thì". Bạn có thể đào tạo máy dự đoán với những ví dụ. Vậy đấy! Các vấn đề mà trước đây không được xem là những vấn đề dự đoán giờ có thể sẽ được giải quyết như vậy. Chúng ta đã thỏa hiệp mà không nhận ra điều đó.

Những sự thỏa hiệp đó là yếu tố chính trong việc đưa ra những quyết định của con người. Người đạt giải Nobel Kinh tế - Herbert Simon gọi đây là "sự tạm chấp nhận". Trong khi những mô hình kinh tế kinh điển dựa trên lí thuyết của những người vô cùng thông minh có thể đưa ra những quyết định lí trí hoàn hảo, Simon nhận ra và nhấn mạnh trong nghiên cứu của ông rằng con người không thể đối mặt với những sự phức tạp. Thay vào đó, họ tạm chấp nhận, làm những gì tốt nhất trong khả năng để đạt được mục tiêu. Suy nghĩ không hề đơn giản, nên con người chọn đường tắt.

Simon là một người am hiểu nhiều lĩnh vực. Ngoài giải thưởng Nobel, ông cũng thắng giải thưởng Turing, thường được gọi là giải thưởng Nobel của máy tính, cho "những sự đóng góp cho trí tuệ nhân tạo". Những sự đóng góp về mặt kinh tế và máy tính đều liên quan với nhau. Lặp lại những suy nghĩ của ông về con người, bài diễn thuyết năm 1976 khi nhận giải thưởng Turing của ông nhấn mạnh rằng máy tính "có nguồn tài nguyên xử lý hạn chế; trong những bước có hạn trong một khoảng thời gian có hạn, chúng có thể thực hiện một số xử lý có hạn." Ông nhận ra rằng máy tính – giống như con người

– cũng tạm chấp nhận[3].

Robot đưa thư và phòng chờ máy bay là ví dụ cho sự tạm chấp nhận khi không có sự dự đoán tốt. Những ví dụ như vậy có ở mọi nơi. Sẽ cần nhiều thực hành và thời gian để tưởng tượng những khả năng có thể đưa ra sự dự đoán tốt hơn. Nó không hoàn toàn là bản năng khi đa số mọi người nghĩ về phòng chờ máy bay như là một giải pháp cho một sự dự đoán không tốt và rằng chúng sẽ không còn nhiều giá trị trong kỷ nguyên của những máy dự đoán mạnh mẽ. Chúng ta đã quen với việc tạm chấp nhận đến mức chúng ta không nghĩ tới những quyết định bao gồm sự dự đoán.

Trong ví dụ về dịch thuật được đề cập trước đó trong cuốn sách, các chuyên gia đánh giá ngôn ngữ dịch thuật tự động không phải là một vấn đề dự đoán mà là một vấn đề về ngôn ngữ. Cách tiếp cận ngôn ngữ truyền thống sử dụng từ điển để dịch từng chữ một, cùng một vài quy tắc ngữ pháp. Đây là sự tạm chấp nhận; nó dẫn đến những kết quả không tốt vì quá nhiều "nếu".

Dịch thuật với máy dự đoán cần sự dự đoán trước câu có nghĩa tương tự trong ngôn ngữ khác. Thống kê số liệu cho phép máy tính lựa chọn bản dịch thuật tốt nhất bằng cách dự đoán "nếu" – một câu mà một người dịch chuyên nghiệp có khả năng sẽ sử dụng dựa vào kết quả tìm kiếm dịch thuật khớp trong dữ liệu. Nó không dựa vào quy tắc ngôn ngữ nào. Người tiên phong trong lĩnh vực này, Frederick Jelinek nhận xét, "Mỗi khi tôi đuổi việc một chuyên gia ngôn ngữ, hiệu suất của máy nhận dạng văn bản tăng lên."[4] Rõ ràng là đây là một sự phát triển đáng gờm với những chuyên gia ngôn ngữ và những dịch giả.

Bằng việc cho phép những quyết định phức tạp hơn, sự dự đoán tốt hơn có thể giảm thiểu rủi ro. Ví dụ, một trong những ứng dụng thực tế gần đây của AI là trong X-quang. Đa phần những chuyên gia X-quang chụp ảnh rồi xác định những vấn đề cần quan tâm. Họ dự đoán sự bất bình thường trong các bức ảnh. AI đang tăng dần khả năng thực hiện chức năng dự đoán ở cấp độ chính xác như con người hoặc hơn, giúp đỡ các chuyên gia X-quang và những chuyên gia y khoa khác trong việc đưa ra quyết định có thể ảnh hưởng đến các bệnh nhân. Chỉ số hiệu suất quan trọng là độ chính xác của sự chẩn đoán: liệu máy có dự đoán bệnh khi bệnh nhân ốm và dự đoán không có bệnh khi bệnh nhân khỏe mạnh không.

Nhưng chúng ta phải xem xét những quyết định này bao gồm những gì. Giả sử các bác sĩ nghi ngờ một khối u và quyết định xác định xem nó có phải ung thư không. Một lựa chọn là chụp ảnh y khoa. Một lựa chọn khác là cái gì đó kĩ lưỡng hơn, ví dụ như sinh thiết. Sinh thiết có lợi thế cung cấp một chẩn đoán chính xác hơn. Vấn đề là sinh thiết khá tốn kém, nên cả bác sĩ và bệnh nhân đều muốn tránh nó nếu vấn đề không nghiêm trọng. Một công việc của chuyên gia X-quang là cung cấp lý do để không tiến hành một thủ tục tốn kém như vậy. Quyết định thực hiện sinh thiết phụ thuộc vào chi phí, mức độ sinh thiết và sẽ tồi tệ thế nào nếu coi nhẹ căn bệnh này. Các bác sĩ sử dụng những yếu tố này để quyết định liệu sinh thiết có đáng về mặt vật lý và chi phí không.

Với sự chẩn đoán đáng tin cậy từ hình ảnh, các bệnh nhân có thể không phải trải qua quá trình sinh thiết kia nữa. Họ không cần phải thỏa hiệp nữa. Những sự tiến bộ trong AI đồng nghĩa với ít nhu cầu đối với sự tạm chấp nhận hơn - nhiều "nếu" và nhiều "thì" hơn. Sự phức tạp hơn đi kèm với ít rủi ro hơn. Điều này làm thay đổi quá trình đưa ra quyết định bằng cách mở rộng nhiều sự lựa chọn hơn.

NHỮNG ĐIỂM CHÍNH

• Sự dự đoán được cải thiện cho phép những người đưa ra quyết định, cho dù là con người hay máy, có thể xử lý được nhiều "nếu-thì" hơn. Điều này dẫn đến những kết quả tốt hơn. Ví dụ, trong trường hợp điều hướng đã được minh họa trong chương này với robot đưa thư, máy dự đoán giải phóng những chiếc xe tự động khỏi sự giới hạn trước đó khi chỉ hoạt động trong những môi trường bị kiểm soát. Những bối cảnh (hoặc trạng thái) này được đặc trưng bởi số lượng "nếu" có hạn. Máy dự đoán cho phép những chiếc xe tự động hoạt động trong những môi trường không cần kiểm soát, ví dụ như trên đường phố, bởi thay vì phải mã hóa trước tất cả những "nếu", máy có thể học cách dự đoán một người điều khiển sẽ làm gì trong bất kỳ tình huống cụ thể. Tương tự như vậy, ví dụ của phòng chờ máy bay minh họa cho việc dự đoán được cải thiện cho phép nhiều "thì" hơn (ví dụ, "rời nhà vào lúc X hoặc Y hoặc Z", phụ thuộc vào dự đoán cần bao nhiêu thời gian để tới sân bay vào một thời điểm cụ thể trong một ngày cụ thể), còn hơn là luôn phải rời đi sớm "đề phòng" và rồi phải dành nhiều thời gian chờ đợi hơn ở phòng chờ sân bay.

• Khi thiếu đi sự dự đoán tốt, chúng ta phải thực hiện nhiều "sự tạm chấp nhận", đưa ra những quyết định "đủ tốt" khi có thông tin có sẵn. Luôn luôn phải rời đi sớm để tới sân bay và thường xuyên phải chờ khi đã tới bởi vì bạn tới quá sớm là một ví dụ của sự tạm chấp nhận. Giải pháp đó không tối ưu, nhưng nó đủ tốt với lượng thông tin có sẵn. Robot đưa thư và phòng chờ sân bay là hai phát minh được thiết kế để đáp ứng sự tạm chấp nhận. Máy dự đoán sẽ giảm thiểu nhu cầu phải tạm chấp nhận và từ đó giảm thiểu lợi nhuận đầu tư vào những giải pháp như hệ thống robot đưa thư và phòng chờ sân bay.

• Chúng ta đã quá quen với việc tạm chấp nhận trong doanh nghiệp và trong đời sống xã hội đến mức sẽ cần phải thực hành để tưởng tượng ra những sự thay đổi khả thi khi máy dự đoán có khả năng xử lý nhiều "nếu" và "thì" hơn, từ đó xử lý nhiều những quyết định phức tạp hơn trong những môi trường phức tạp hơn. Việc đa số mọi người nghĩ tới phòng chờ sân bay như là một giải pháp cho sự dự đoán không tốt không hoàn toàn là bản năng và chúng sẽ ít có giá trị hơn trong kỷ nguyên của những máy dự đoán mạnh mẽ. Một ví dụ khác là việc sử dụng sinh thiết, tồn tại vì những yếu điểm trong dự đoán từ những hình ảnh y khoa. Khi sự tự tin của máy dự đoán tăng lên, ảnh hưởng đến hình ảnh y khoa của AI cũng sẽ lớn hơn nhiều lên những công việc liên quan đến việc thực hiện sinh thiết vì, cũng như phòng chờ sân bay, thủ tục tốn kém và đòi hỏi này được phát minh vì sự dự đoán không tốt. Máy dự đoán sẽ cung cấp những phương pháp mới và tốt hơn để quản lý rủi ro.

11Những quyết định được tự động hóa hoàn toàn

V

ào ngày 12 tháng 12 năm 2016, thành viên “jmdavis” của câu lạc bộ Tesla Motors đã đăng lên diễn đàn xe điện để nói về trải nghiệm anh có với chiếc Tesla của mình. Khi đang lái xe đi làm trên đường cao tốc Florida ở tốc độ khoảng 60 dặm một giờ, bảng điều khiển Tesla chỉ ra một chiếc xe hơi phía trước mà anh ta không thể nhìn thấy vì chiếc xe tải ngay phía trước đã chặn tầm nhìn. Đột nhiên, phanh khẩn cấp hoạt động, cho dù chiếc xe tải phía trước chưa dừng lại. Một giây sau, chiếc xe tải tạt vào lề đường để tránh đụng phải chiếc xe đã dừng đột ngột vì những mảnh vụn trên đường ở phía trước. Chiếc Tesla quyết định phanh lại nhanh chóng trước khi chiếc xe tải kịp làm như vậy, cho phép chiếc xe của jmdavis dừng lại với khoảng cách an toàn. Anh viết:

Nếu tôi tự lái xe lúc đó, có thể tôi sẽ không dừng kịp thời, bởi vì tôi không thể nhìn thấy chiếc xe phía trước đã dừng. Chiếc xe phản ứng lại thậm chí trước khi chiếc xe phía trước tôi phản ứng và điều đó làm nên sự khác biệt giữa một vụ đụng xe và một cú phanh kịp. Làm tốt lắm Tesla, cảm ơn vì đã cứu mạng tôi.[1]

Tesla đã gửi bản cập nhật phần mềm tới xe của nó để cho phép chức năng tự lái Autopilot của nó khám phá ra thông tin ra-đa để có được bức ảnh rõ hơn về không gian xung quanh xe.[2] Khi chức năng của Tesla hoạt động khi xe ở trong chế độ tự lái, có thể dễ dàng tưởng tượng một tình huống mà chiếc xe nắm quyền kiểm soát khi tai nạn sắp xảy ra. Những nhà sản xuất xe ở Hoa Kỳ đã đạt được thỏa thuận với Bộ Giao thông vận tải để tạo ra phanh khẩn cấp tự động theo chuẩn trong những phương tiện cho tới năm 2022.[3] Thông thường, sự khác biệt giữa AI và máy tự động hóa không rõ ràng. Sự tự động hóa phát sinh khi một máy thực hiện toàn bộ công việc, chứ không phải là dự đoán. Ở thời điểm viết, con người vẫn cần can thiệp định kỳ vào việc lái xe. Khi nào chúng ta mới có một sự tự động hóa hoàn toàn?

AI, trong hiện thân hiện tại của nó, liên quan đến một yếu tố: dự đoán. Mỗi yếu tố thể hiện sự bổ sung cho dự đoán, điều trở nên có ích khi sự dự đoán trở nên rẻ hơn. Liệu sự tự động hóa hoàn toàn có hợp lý hay không phụ thuộc vào việc máy cũng thực hiện những yếu tố khác.

Con người và máy có thể tích lũy dữ liệu, cho dù là đầu vào, đào tạo, hay phản hồi, phụ thuộc vào loại dữ liệu. Con người cuối cùng phải đưa ra sự đánh giá, nhưng con người có thể mã hóa sự đánh giá và lập trình nó vào máy trước khi có sự dự đoán. Hoặc máy có thể học cách dự đoán sự đánh giá của con người thông qua phản hồi. Điều này khiến chúng ta hành động. Khi nào thì máy sẽ thay thế con người thực hiện hành động? Tinh tế hơn, từ khi nào việc máy xử lý dự đoán gia tăng lợi ích cho máy hơn là khi con người thực hiện cùng một hành động? Chúng ta phải xác định các phản hồi để máy thực hiện những yếu tố khác (thu thập dữ liệu, đánh giá, hành động), từ đó quyết định liệu một công việc nên hoặc sẽ được tự động hóa hoàn toàn.

Kính râm trong đêm

Vùng Pilbara hẻo lánh của nước Úc có nhiều quặng sắt. Đa số những mỏ khai thác đều cách khoảng 100 dặm từ thành phố lớn gần nhất, Perth. Tất cả các công nhân ở mỏ khai thác đều làm nhiều ca tăng cường kéo dài nhiều tuần. Họ nhận được mức lương tốt tương ứng và được hỗ trợ khi ở mỏ khai thác. Không ngạc nhiên khi các công ty khai thác mỏ muốn tận dụng công nhân khi họ ở đó.

Các mỏ quặng sắt lớn ở mỏ khai thác Rio Tinto khổng lồ trải dài và có giá trị lớn về kinh tế. Họ lấy quặng sắt từ phía trên mặt đất trong những hố khổng lồ đến mức một vụ va chạm thiên thạch cũng không thể sánh bằng. Do đó, công việc chính của các công nhân là sử dụng những chiếc xe tải có kích thước của một ngôi nhà hai tầng, không chỉ đi lên từ hố mà còn tới những tuyến đường sắt xây dựng gần đó để vận chuyển quặng sắt hàng nghìn dặm về phía Bắc tới cảng đợi. Chi phí thực cho các công ty khai thác mỏ không phải là ở con người mà là ở thời gian chết. Các công ty khai thác mỏ đã cố gắng tối ưu hóa bằng việc chạy suốt đêm. Tuy nhiên, ngay cả những người làm ca dài nhất cũng không thể hoạt động hiệu quả vào ban đêm. Ban đầu, Rio Tinto giải quyết một vài vấn đề khai thác con người bằng việc sử dụng nhiều xe tải có thể điều khiển từ Perth.[4] Nhưng vào năm 2016, họ đã có một bước tiến lớn

với 73 chiếc xe tải tự lái.[5] Sự tự động hóa này đã tiết kiệm cho Rio Tinto 15% chi phí vận hành. Mỏ khai thác cho xe chạy 24 giờ mỗi ngày mà không có thời gian nghỉ và không có điều hòa trong xe cho dù nhiệt độ lên ngưỡng 55oC vào ban ngày. Cuối cùng, vì không cần người lái, những chiếc xe tải không cần mặt trước và mặt sau, nghĩa là chúng không cần phải quay lại, nhờ đó tiết kiệm hơn về mặt an toàn, không gian, sự bảo trì và tốc độ.

AI đã thực hiện được điều này bằng cách dự đoán những rủi ro trên đường đi của xe tải và điều phối những đường lái chuyên biệt. Không cần người lái xe nào quan sát sự an toàn của xe tại chỗ hay từ xa. Và cần ít người để đảm bảo an toàn hơn. Tiến xa hơn nữa, những công ty khai thác mỏ ở Canada đang nghiên cứu để mang đến những robot AI có thể đào dưới mặt đất, trong khi những công ty ở Úc đang tìm cách tự động hóa toàn bộ chuỗi từ mặt đất đến cảng (bao gồm máy đào đất, xe ủi đất và tàu). Khai thác mỏ là cơ hội tuyệt vời cho sự tự động hóa hoàn toàn bởi vì nó đã loại bỏ con người khỏi nhiều hoạt động. Ngày nay, con người thực hiện những chức năng chỉ đạo chính nhưng quan trọng. Trước khi có những tiến bộ gần đây trong AI, mọi thứ ngoại trừ sự dự đoán đều có thể được tự động hóa. Máy dự đoán thực hiện bước cuối cùng trong việc loại bỏ con người khỏi những công việc liên quan. Trước đó, con người tìm hiểu môi trường xung quanh và nói cho thiết bị cần phải làm gì. Hiện giờ, AI thu thập thông tin từ những máy cảm biến có thể học cách dự đoán những trở ngại trên đường. Bởi vì máy dự đoán có thể dự đoán liệu con đường có bị cản trở hay không, các công ty khai thác mỏ không còn cần con người làm vậy nữa.

Nếu yếu tố con người cuối cùng trong một công việc là sự dự đoán, thì một khi máy dự đoán có thể làm tốt ngang với con người, người đưa ra quyết định có thể loại bỏ con người hoàn toàn khỏi phép tính. Tuy nhiên, như chúng ta sẽ thấy trong chương này, có rất ít những công việc rõ ràng như khai thác mỏ. Với đa số những quyết định tự động hóa, sự cung cấp dự đoán của máy không nhất thiết nghĩa là phải loại bỏ sự đánh giá của con người và thay thế bằng máy đưa ra quyết định, hay loại bỏ hành động của con người và thay thế bằng một robot vật lý.

Không có thời gian hay nhu cầu suy nghĩ

Máy dự đoán khiến những chiếc xe tự lái như Tesla trở nên khả thi. Nhưng sử

dụng máy dự đoán để kích hoạt máy giành sự điều khiển xe từ tay con người lại là một vấn đề khác. Tỉ lệ ở đây rất dễ hiểu: giữa thời điểm một tai nạn được dự đoán và phản ứng cần thiết, con người không có thời gian để suy nghĩ hoặc hành động ("không có thời gian để suy nghĩ"). Ngược lại, tương đối dễ dàng để có thể lập trình phản ứng của xe. Khi tốc độ là cần thiết, lợi ích của việc nhường lại quyền kiểm soát cho máy móc là rất cao.

Khi bạn sử dụng một máy dự đoán, sự dự đoán phải được thông báo cho người đưa ra quyết định. Nhưng nếu sự dự đoán trực tiếp dẫn đến một hành động hiển nhiên ("không cần suy nghĩ"), thì trường hợp bỏ qua sự đánh giá của con người trong chuỗi lại giảm đi. Nếu máy có thể được mã hóa để đánh giá và xử lý hành động hệ quả tương đối dễ dàng, thì để máy xử lý toàn bộ công việc là hợp lý.

Điều này dẫn đến các cách thức đổi mới. Ở Thế vận hội Olympic Rio 2016, một chiếc máy ghi hình mới đã quay các vận động viên bơi lội dưới nước bằng việc theo dõi hành động và chuyển động để ghi hình được từ dưới bể.[6] Trước đó, người ta điều khiển máy ghi hình từ xa nhưng phải dự đoán vị trí của người bơi. Bây giờ, máy dự đoán có thể làm được điều đó. Việc bơi lội chỉ là khởi đầu mà thôi. Các chuyên gia đang nghiên cứu việc ứng dụng tính tự động của máy ghi hình vào những môn thể thao phức tạp hơn như bóng rổ.[7] Một lần nữa, nhu cầu về tốc độ và khả năng mã hóa sự đánh giá đang thúc đẩy trạng thái tự động hóa hoàn toàn.

Sự phòng ngừa tai nạn và những máy ghi hình thể thao tự động có điểm gì chung? Ở cả hai trường hợp, hành động nhanh dựa trên sự dự đoán có thể đem lại nhiều lợi ích và sự đánh giá có thể được mã hóa hoặc dễ dàng dự đoán. Sự tự động hóa xảy ra khi việc để máy xử lý tất cả các chức năng đem lại nhiều lợi ích hơn việc bao gồm con người trong quá trình.

Sự tự động hóa cũng có thể phát sinh khi việc giao tiếp trở nên tốn kém. Khám phá vũ trụ chẳng hạn. Việc gửi robot vào vũ trụ dễ hơn việc gửi con người. Nhiều công ty đang nghiên cứu cách khai thác những khoáng sản quý giá từ Mặt Trăng, nhưng họ cần vượt qua nhiều thách thức kỹ thuật. Điều khiến chúng ta lo ngại nhất ở đây là những robot ở Mặt Trăng sẽ điều hướng và hành động ra sao. Mất ít nhất hai giây để tín hiệu radio lên tới Mặt Trăng và ngược lại, vậy nên vận hành robot trên Mặt Trăng từ Trái Đất là một quá

trình chậm chạp và vất vả. Như vậy robot không thể phản ứng nhanh trong những tình huống mới. Nếu một con robot di chuyển dọc theo bề mặt của Mặt Trăng đột nhiên gặp phải một mỏm đá, bất kỳ sự giao tiếp chậm trễ nào cũng sẽ đồng nghĩa với việc hướng dẫn từ mặt đất có thể đến quá muộn. Máy dự đoán cung cấp một giải pháp. Với sự dự đoán tốt, những hành động của robot trên Mặt Trăng có thể được tự động hóa mà không cần con người dưới mặt đất hướng dẫn từng bước. Không có AI, những sự đầu tư thương mại như vậy khó có thể xảy ra.

Khi luật pháp yêu cầu con người hành động

Khái niệm rằng sự tự động hóa có thể dẫn đến rủi ro đã trở thành một chủ đề phổ biến trong khoa học viễn tưởng. Ngay cả khi chúng ta đều thoải mái với việc tự động hóa máy hoàn toàn, luật pháp có thể sẽ không cho phép như vậy. Isaac Asimov* đã đưa ra ba điều luật cho robot khó mã hóa, được thiết kế khéo léo để loại bỏ khả năng robot có thể gây hại cho con người.[8]

* *Isaac Asimov là một tác giả người Mỹ và là giáo sư hóa sinh tại Đại học Boston. Ông nổi tiếng nhất với các tác phẩm về khoa học viễn tưởng.*

Tương tự, những nhà triết học hiện đại thường đặt ra những tình huống khó xử về mặt đạo đức mà dường như có vẻ trừu tượng. Hãy xem xét tình huống xe đẩy: Hãy tưởng tượng rằng mình đang đứng ở một nơi cho phép bạn đẩy một chiếc xe từ đoạn đường này sang đoạn đường khác. Bạn thấy có năm người đang trên đường xe đẩy của bạn. Bạn có thể đổi sang đoạn đường khác, nhưng có một người đang đi trên đoạn đường đó. Bạn không có lựa chọn nào khác và không có thời gian để suy nghĩ. Bạn sẽ làm gì? Câu hỏi đó gây bối rối nhiều người, và thường họ sẽ chỉ muốn tránh suy nghĩ về những câu hỏi hóc búa. Tuy nhiên, với những chiếc xe tự lái, tình huống đó có thể xảy ra. Một ai đó sẽ phải giải quyết vấn đề nan giải và lập trình phản ứng phù hợp cho xe. Vấn đề này không thể tránh khỏi. Ai đó – khả năng cao là luật pháp – sẽ quyết định ai sống và ai chết.

Hiện tại, thay vì mã hóa những lựa chọn về mặt đạo đức của chúng ta vào máy tự động, chúng ta chọn cách giữ con người lại trong chuỗi. Ví dụ, tưởng tượng một vũ khí bay có thể hoạt động hoàn toàn tự động – xác định, nhắm mục tiêu và giết kẻ thù. Ngay cả khi một sỹ quan quân đội có thể tìm được

một máy dự đoán có thể phân biệt dân thường với lính, sẽ mất bao lâu để các chiến sĩ tìm ra cách gây bối rối cho máy dự đoán? Mức độ yêu cầu chính xác không phải lúc nào cũng có sẵn. Vì vậy vào năm 2012, Bộ Quốc phòng Hoa Kỳ đã đưa ra một chỉ thị mà nhiều người hiểu là sự bắt buộc giữ con người trong chuỗi những quyết định có nên tấn công hay không.[9] Tuy không rõ phải thực hiện hành động đó thường xuyên hay không, nhu cầu cần sự can thiệp của con người, với bất kỳ lý do nào, cũng sẽ hạn chế quyền tự chủ của máy dự đoán ngay cả khi chúng có thể tự hoạt động.[10] Ngay cả phần mềm Autopilot của Tesla – cho dù có thể tự lái xe – vẫn đi kèm với những điều khoản pháp lý và điều kiện yêu cầu người lái xe luôn đặt tay lên bánh lái. Theo quan điểm của chuyên gia kinh tế, điều này có hợp lý hay không còn phụ thuộc vào bối cảnh của rủi ro có thể xảy ra. Ví dụ ,vận hành một chiếc xe tự động trong một mỏ khai thác từ xa hoặc một tầng của nhà máy có thể khá khác với việc vận hành trên những con đường công cộng. Điều khiến môi trường "bên trong nhà máy" khác biệt với "đường công cộng" là khả năng của điều mà các chuyên gia kinh tế gọi là "ngoại tác" – những điều mà người khác phải gánh chịu, thay vì những người đưa ra quyết định quan trọng.

Các chuyên gia kinh tế có các giải pháp khác nhau cho vấn đề ngoại tác. Một giải pháp là phân công trách nhiệm để người đưa ra quyết định chính tiếp thu những chi phí bên ngoài khác. Nhưng khi nói về máy tự động, việc nhận dạng bên liên quan rất phức tạp. Máy càng gần với những rủi ro tiềm ẩn bên ngoài tổ chức (và, đương nhiên, gần với những rủi ro của con người trong tổ chức), thì khả năng cao là nó sẽ thận trọng và yêu cầu giữ con người lại trong chuỗi một cách hợp pháp.

Khi con người giỏi hơn trong việc hành động

Câu hỏi: Cái gì màu cam và nghe giống từ con vẹt (parrot)?

Đáp án? Củ cà rốt (carrot).

Trò đùa đó có thú vị không? Hay là trò đùa này: Một cô gái nhỏ hỏi bố: "Bố? Có phải tất cả những câu chuyện cổ tích đều bắt đầu bằng 'ngày xửa ngày xưa' không?" Người bố trả lời: "Không, có rất nhiều câu chuyện cổ tích bắt đầu bằng 'Nếu đắc cử, tôi hứa...'"

Đồng ý là các chuyên gia kinh tế không phải là những người kể chuyện cười

giỏi nhất. Nhưng chúng tôi giỏi hơn so với máy móc. Chuyên gia nghiên cứu Mike Yeomans và đồng tác giả phát hiện ra rằng nếu con người nghĩ rằng máy gợi ý một trò đùa, họ sẽ thấy nó ít thú vị hơn là khi con người gợi ý. Các chuyên gia nghiên cứu thấy rằng máy có thể làm tốt việc gợi ý trò đùa, nhưng con người thích tin vào những gợi ý của con người hơn. Con người đọc những trò đùa sẽ thấy hài lòng nhất nếu được nói là những gợi ý đó đến từ con người, nhưng thực chất những gợi ý được quyết định bởi máy.

Điều này cũng đúng trong những thành tựu nghệ thuật và thi đấu thể thao. Sức mạnh của nghệ thuật thường xuất phát từ kiến thức của nhà tài trợ về trải nghiệm con người của người nghệ sĩ. Một phần của sự hồi hộp khi theo dõi một sự kiện thể thao phụ thuộc vào việc con người đang thi đấu. Ngay cả khi máy có thể chạy nhanh hơn con người, kết quả của cuộc thi cũng sẽ ít thú vị hơn. Chơi với trẻ em, quan tâm tới người già và nhiều hành động khác liên quan đến tương tác xã hội có thể cũng sẽ tốt hơn khi con người thực hiện. Ngay cả khi máy biết trình bày thông tin cho một đứa trẻ với mục đích giáo dục; đôi khi sẽ tốt hơn khi con người truyền tải những thông tin đó. Có lẽ theo thời gian, chúng ta có thể chấp nhận việc robot chăm sóc chúng ta và con cái của chúng ta hơn và chúng ta có thể thậm chí thấy thích thú khi theo dõi những cuộc thi thể thao robot, nhưng hiện tại, con người vẫn thích có một số hành động được thực hiện bởi những người khác hơn.

Khi con người phù hợp nhất để thực hiện hành động, những quyết định đó sẽ không hoàn toàn được tự động hóa. Vào những thời điểm khác, sự dự đoán là ràng buộc chính của sự tự động hóa. Khi sự dự đoán đủ tốt và đánh giá những sự trả giá có thể được xác định trước – cho dù là con người tiến hành mã hóa hay máy học thông qua quan sát con người – thì quyết định vẫn sẽ được tự động hóa.

NHỮNG ĐIỂM CHÍNH

• Việc đưa AI vào một công việc không thực sự ám chỉ việc tự động hóa hoàn toàn công việc đó. Sự dự đoán chỉ là một thành phần. Trong nhiều trường hợp, con người vẫn cần phải áp dụng sự đánh giá và thực hiện hành động. Tuy nhiên, đôi khi sự đánh giá có thể khó mã hóa hoặc nếu có sẵn đủ ví dụ, máy có thể học cách để dự đoán sự đánh giá. Bên cạnh đó, máy có thể thực hiện hành động. Khi máy thực hiện tất cả các yếu tố của một công việc, công việc khi đó sẽ được tự động hóa hoàn toàn và con người sẽ hoàn toàn bị loại

bỏ khỏi chuỗi.

• Công việc có khả năng tự động hóa hoàn toàn đầu tiên là những công việc mà ở đó sự tự động hóa hoàn toàn mang lại lợi nhuận cao nhất. Đó bao gồm những công việc mà: (1) những yếu tố khác đã được tự động hóa ngoại trừ sự dự đoán (ví dụ, khai thác mỏ); (2) lợi ích của việc phản ứng nhanh khi hành động là rất cao (ví dụ, xe không người lái); (3) lợi nhuận của việc giảm thời gian chờ đợi sự dự đoán là rất cao (ví dụ, khám phá vũ trụ).

• Điểm khác biệt quan trọng giữa xe tự động hoạt động trên đường phố so với những xe hoạt động ở mỏ khai thác là xe hoạt động trên phố tạo ra những ngoại tác đáng kể trong khi xe hoạt động ở mỏ khai thác thì không. Xe tự động hoạt động trên đường phố có thể gây ra tai nạn, phát sinh chi phí cho những cá nhân không đưa ra quyết định. Ngược lại, những tai nạn gây ra bởi xe tự động hoạt động tại mỏ khai thác chỉ phát sinh chi phí ảnh hưởng đến những tài sản hoặc những ai liên quan đến mỏ khai thác. Chính phủ điều phối những hoạt động gây ra những ngoại tác. Do vậy, sự điều phối là rào cản tiềm năng cho sự tự động hóa hoàn toàn của những ứng dụng tạo ra những ngoại tác đáng kể. Sự phân công trách nhiệm là một công cụ phổ biến được sử dụng bởi các chuyên gia kinh tế để giải quyết vấn đề này bằng việc tiếp thu những ngoại tác. Chúng tôi mong chờ một làn sóng thay đổi chính sách phát triển liên quan đến sự phân công trách nhiệm được thúc đẩy bởi nhu cầu gia tăng trong nhiều lĩnh vực mới của sự tự động hóa.

PHẦN 3CÔNG CỤ



12Tái xây dựng luồng công việc

G

iữa cuộc cách mạng IT, các doanh nghiệp đã đặt ra câu hỏi: “Chúng ta nên áp dụng máy tính vào doanh nghiệp như thế nào?” Với một số người, câu trả lời rất đơn giản: “Tìm nơi mà chúng ta thực hiện nhiều sự tính toán và thay thế con người bởi máy tính; chúng tốt hơn; nhanh hơn và giá thành rẻ hơn.” Với những doanh nghiệp khác, điều này ít rõ ràng hơn. Tuy nhiên, họ đã thử nghiệm. Nhưng kết quả của những thử nghiệm đó cần thời gian để thực hiện. Robert Solow, một chuyên gia kinh tế đạt giải Nobel, than thở rằng, “Bạn có thể nhìn thấy thời đại máy tính ở khắp mọi nơi trừ năng suất thông kê số liệu.”[1]

Từ thử thách này dẫn đến một phong trào kinh doanh thú vị gọi là “tái cấu trúc”. Vào năm 1993, Michael Hammer và James Champy đã lập luận trong cuốn sách *Reengineering the Corporation* (tạm dịch: Tái xây dựng các tập đoàn) của họ rằng, để sử dụng công nghệ đa năng mới – máy tính – các doanh nghiệp cần phải xem lại các quy trình của họ và vạch ra mục tiêu họ muốn đạt được. Các doanh nghiệp sau đó cần nghiên cứu luồng công việc của họ và xác định những công việc nào cần thực hiện để đạt được mục tiêu và chỉ khi đó mới xem xét liệu máy tính có đóng vai trò gì trong những công việc này không.

Một trong những ví dụ yêu thích của Hammer và Champy là sự tiến thoái lưỡng nan mà Ford phải đối mặt vào những năm 1980, không phải với việc sản xuất xe mà với việc trả tiền cho nhân viên.[2] Ở chi nhánh Bắc Hoa Kỳ, tài khoản cần thanh toán tại các bộ phận lên đến 500 nhân viên, và Ford hy vọng rằng việc đầu tư vào máy tính có thể giảm con số này tới 20%. Mục tiêu chỉ có 400 người ở các bộ phận đó là không thực tế; vì đối thủ cạnh tranh của Ford - Mazda chỉ có năm người trong số tài khoản cần thanh toán.

Để đạt được hiệu suất cao hơn, những người quản lý ở Ford đã xem xét lại quy trình giao dịch. Giữa thời gian đơn hàng giao dịch được viết và thực sự được phát hành để mua cái gì đó, có rất nhiều nhân viên xử lý đơn. Nếu chỉ một trong số những người này tốn nhiều thời gian để làm việc, thì toàn bộ hệ

thống sẽ bị trì trệ. Giả sử, nếu chỉ một lượng nhỏ đơn hàng có vấn đề, thì đa số thời gian của người đó sẽ dành vào việc giải quyết chúng. Điều đó khiến mỗi đơn hàng được giải quyết với tốc độ rất chậm.

Việc này chứa đựng tiềm năng để sử dụng máy tính hiệu quả. Không những máy tính có thể giảm thiểu sự kết hợp không phù hợp, làm chậm trễ hệ thống, mà nó còn có thể phân loại những trường hợp khó với những trường hợp dễ hơn và đảm bảo những trường hợp dễ hơn được xử lý với tốc độ hợp lý. Một khi hệ thống mới hoạt động, bộ phận số tài khoản cần thanh toán của Ford giảm đến 75%, toàn bộ quy trình trở nên nhanh hơn và chính xác hơn một cách đáng kể.

Không phải trường hợp tái cấu trúc nào cũng liên quan đến việc giảm số lượng đầu người, cho dù nhiều người nghĩ về nó đầu tiên.[3] Tổng quát hơn, sự tái cấu trúc có thể cải thiện chất lượng dịch vụ. Một ví dụ khác, Mutual Benefit Life, một công ty bảo hiểm nhân thọ lớn, nhận ra rằng trong khi xử lý những ứng dụng, 19 người ở năm bộ phận thực hiện 30 bước khác nhau. Nếu bạn dùng một ứng dụng đặc thù vào quá trình đó, bạn có thể hoàn thành nó chỉ trong một ngày. Nhưng thay vào đó, ứng dụng này sẽ mất từ 5 đến 25 ngày. Vì sao? Thời gian vận chuyển. Tệ hơn là, nhiều sự thiếu hiệu quả khác chồng chất lên bởi vì họ có thể quá tập trung vào một mục tiêu chậm. Một lần nữa, hệ thống cơ sở dữ liệu chung của một hệ thống máy tính doanh nghiệp cải thiện quá trình đưa ra quyết định, giảm thiểu sự xử lý và cải thiện năng suất đáng kể. Cuối cùng, một người có thẩm quyền sử dụng ứng dụng này, có thể xử lý công việc trong khoảng bốn tiếng đến một vài ngày.

Giống như máy tính cổ điển, AI là một công nghệ đa năng. Nó có tiềm năng để ảnh hưởng lên mọi quyết định, bởi vì sự dự đoán là thông tin đầu vào quan trọng của quá trình đưa ra quyết định. Do vậy, không có người quản lý nào sẽ đạt được năng suất cao hơn chỉ nhờ “thực hiện một vài AI” ở một vấn đề hoặc vào trong một quy trình hiện có. Thay vào đó, AI là loại hình công nghệ đòi hỏi sự xem xét lại quy trình giống như cách mà Hammer và Champy đã làm. Các doanh nghiệp đang tiến hành phân tích về những luồng công việc và chia chúng thành nhiều thành phần công việc. Giám đốc tài chính của Goldman, Sachs R. Martin Chavez nhận xét rằng, 146 công việc khác nhau trong quá trình chào bán công khai đầu tiên đã “cầu mong được tự động hoá”.[4] Nhiều trong số 146 công việc đó được dự đoán với quyết định rằng

các công cụ AI sẽ được tăng cường đáng kể. Khi ai đó viết về sự thay đổi của Goldman Sachs một thập kỷ trước, đa phần các câu chuyện là về vai trò quan trọng của AI trong sự chuyển đổi đó.

Việc thực hiện AI trên thực tế là thông qua sự phát triển của các công cụ. Đơn vị thiết kế công cụ AI không phải là "nhiệm vụ" hay "nghề nghiệp" hay "chiến lược", mà là "công việc". Các công việc là tập hợp của những sự quyết định (giống như những quyết định được thể hiện ở hình 7-1 và được phân tích trong phần hai). Những quyết định dựa trên sự dự đoán và đánh giá và được thông báo bởi dữ liệu. Sự khác nhau của chúng nằm ở hành động theo sau đó (Xem hình 12-1).



Đôi khi chúng ta có thể tự động hoá tất cả các quyết định chỉ trong một nhiệm vụ. Hoặc chúng ta có thể tự động hoá những quyết định cuối cùng chưa được tự động hoá bởi vì sự dự đoán đã được cải thiện. Sự gia tăng của các máy dự đoán thúc đẩy suy nghĩ về việc làm thế nào để thiết kế lại và tự động hoá toàn bộ các quy trình, hoặc điều mà chúng tôi gọi là "các luồng công việc", hiệu quả loại bỏ con người khỏi những công việc đó. Nhưng vì sự dự đoán tốt hơn và có giá thành rẻ hơn đã dẫn đến sự tự động hoá thuần tuý, việc sử dụng máy dự đoán cũng cần phải đem lại nhiều lợi ích khi áp dụng vào giải quyết các nhiệm vụ. Nếu không bạn sẽ muốn sử dụng một máy dự đoán làm việc dưới sự giám sát của những người đưa ra quyết định.

Ảnh hưởng của các công cụ AI lên các luồng công việc

Cho đến nay, chúng tôi đã chứng kiến hơn 150 công ty AI trong CDL, phòng thí nghiệm giúp đỡ những công ty nền tảng khoa học phát triển của chúng tôi. Mỗi công ty tập trung phát triển một công cụ AI có thể giải quyết một công việc cụ thể trong luồng công việc cụ thể. Một công ty khởi nghiệp dự đoán những đoạn văn quan trọng nhất trong một tài liệu và làm nổi bật chúng. Một công ty khác dự đoán những lỗi sản xuất và đánh dấu chúng. Một công ty khác dự đoán những phản hồi dịch vụ khách hàng thích hợp và trả lời các câu hỏi. Danh sách vẫn còn dài. Những công ty lớn đang bổ sung hàng trăm nếu không phải hàng nghìn AI khác nhau để nâng cao những công việc khác nhau trong các luồng công việc. Thực tế, Google đang phát triển hàng trăm công cụ AI khác nhau để giải quyết các công việc đa dạng, từ email đến dịch thuật

hoặc lái xe.[5] Với nhiều doanh nghiệp, máy dự đoán có thể sẽ có ảnh hưởng, nhưng là theo cách tăng dần và đa phần là không dễ thấy, giống như cách mà AI cải thiện nhiều ứng dụng ảnh trên điện thoại thông minh của bạn. Nó phân loại các bức ảnh theo cách hữu ích nhưng không làm ảnh hưởng lớn đến cách bạn sử dụng ứng dụng.

Tuy nhiên, có thể bạn đang đọc cuốn sách này vì bạn quan tâm đến cách mà AI mang đến những thay đổi quan trọng trong doanh nghiệp của bạn. Các công cụ AI có thể thay đổi các luồng công việc theo hai cách. Đầu tiên, chúng có thể khiến các công việc trở nên lỗi thời và do đó loại bỏ chúng khỏi các luồng công việc. Thứ hai, chúng có thể thêm vào những công việc mới. Điều này có thể khác với từng doanh nghiệp và với từng luồng công việc.

Hãy cùng xem xét vấn đề của việc tuyển sinh viên vào chương trình MBA, một quy trình mà chúng ta đã quá quen thuộc. Bạn có thể đã từng ở bên này hoặc bên kia của các quy trình tuyển dụng tương tự, có lẽ để tuyển dụng nhân viên hoặc để khiến các khách hàng đăng ký. Các luồng công việc tuyển dụng MBA bắt đầu với nhiều hồ sơ tiềm năng và dẫn đến những nhóm người nhận được và chấp nhận lời mời gia nhập. Nó có ba phần rộng: (1) một kênh bán hàng bao gồm các bước được thiết kế để hỗ trợ việc nộp hồ sơ, (2) một quy trình cân nhắc xem ai sẽ được nhận lời mời, và (3) những bước thúc đẩy những người nhận được lời mời chấp nhận chúng. Mỗi phần đều cần sự phân bổ nguồn lực đáng kể.

Rõ ràng mục tiêu của bất kỳ quy trình tuyển sinh nào là để đạt được một khóa có nhiều sinh viên giỏi nhất. Tuy nhiên, điều gì “tốt nhất” lại là một câu hỏi phức tạp và cũng liên quan đến những mục tiêu chiến lược của trường. Chúng tôi sẽ tạm gác lại việc những định nghĩa khác nhau của “điều tốt nhất” có ảnh hưởng thế nào lên việc thiết kế những công cụ AI cũng như những nhiệm vụ trong các luồng công việc, và đơn giản giả định rằng các trường có định nghĩa rõ ràng về điều gì là tốt nhất cho tổ chức. Trên thực tế, bước trung gian trong luồng công việc tuyển dụng – lựa chọn ứng viên để phỏng vấn – bao gồm những quyết định quan trọng liên quan đến liệu các lời mời phỏng vấn nên được đưa ra sớm hay muộn trong quy trình và liệu chúng có đi kèm những mục đích liên quan đến tài chính hoặc sự hỗ trợ đi kèm. Những quyết định đó còn hơn cả việc lựa chọn những người tốt nhất đơn thuần nhưng cũng dự đoán được phương pháp hiệu quả nhất để những người tốt nhất chấp nhận

lời mời (đôi khi điều này xảy ra sau trong luồng công việc).

Những hệ thống xếp hạng hồ sơ hiện tại bao gồm những sự đánh giá thô. Các ứng viên thường được xếp vào các nhóm a, b và c, trong đó (a) chắc chắn sẽ nhận được lời mời; (b) nên nhận được lời mời nếu (a) từ chối lời mời; và (c) không nhận được lời mời. Điều đó dẫn đến nhu cầu quản lý rủi ro để cân bằng ưu và nhược điểm của các hành động có thể gia tăng khả năng sai sót. Ví dụ, bạn không muốn đặt ai đó vào nhóm (c) trong khi đáng lẽ họ nên ở nhóm (a) hoặc thậm chí là (b) vì những lý do không rõ ràng trong hồ sơ. Tương tự, bạn không muốn đặt ai đó vào nhóm (a) trong khi đáng lẽ họ nên xếp thấp hơn. Vì hồ sơ có tính đa chiều, nên các đánh giá có thể khiến các ứng viên được xếp vào các nhóm là sự kết hợp của ý kiến chủ quan và khách quan.

Giả sử chương trình MBA phát triển AI có khả năng nhận hồ sơ và những thông tin khác – có lẽ là những video phỏng vấn mà mọi người thường nộp, cùng với những thông tin công khai có sẵn được đăng trên phương tiện truyền thông – đồng thời đào tạo AI ghi nhận những thông tin quá khứ từ đơn ứng tuyển và thông tin điểm để cung cấp những thứ hạng rõ ràng cho ứng cử viên. Công cụ AI sẽ khiến công việc lựa chọn ứng viên nên nhận được lời mời nhanh hơn, có giá thành rẻ hơn và chính xác hơn. Câu hỏi quan trọng là: Công nghệ dự đoán kỳ diệu như vậy có ảnh hưởng như thế nào đến phần còn lại của luồng công việc MBA?

Công nghệ giả định xếp thứ hạng các ứng viên của chúng tôi cung cấp một sự dự đoán cho biết những ứng viên nào có khả năng là tốt nhất. Điều này sẽ ảnh hưởng tới những quyết định trong suốt luồng công việc. Chúng bao gồm những lời mời sớm (có lẽ trước những trường khác), những động lực về tài chính (học bổng), và sự quan tâm đặc biệt (những bữa ăn trưa với giảng viên hoặc những sinh viên nổi bật). Đây đều là những quyết định có sự trả giá và nguồn tài nguyên hiếm có. Việc có một danh sách chính xác những ứng viên mong muốn sẽ thay đổi việc ai nhận được những nguồn tài nguyên này. Chúng ta cũng có thể sẵn lòng chi trả nhiều động lực về tài chính hơn cho những ứng viên mà chúng ta tin là tốt nhất.

Sự xếp hạng dự đoán có thể sẽ có ảnh hưởng lớn lên những quyết định được thực hiện trước khi trường nhận được các hồ sơ. Nhiều trường biết rằng trong khi họ muốn nhận được nhiều hồ sơ hơn, nếu họ nhận được quá nhiều, họ sẽ

đối mặt với vấn đề đánh giá và xếp hạng chúng. Máy dự đoán của chúng tôi giảm thiểu chi phí việc thực hiện những xếp hạng như vậy. Kết quả là nó sẽ làm tăng lợi nhuận của việc có nhiều hồ sơ hơn để xếp hạng. Điều này đặc biệt đúng nếu công nghệ cũng có thể đánh giá mức độ nghiêm túc của hồ sơ (vì nó rất tuyệt, tại sao không?). Do đó, trường học có thể mở rộng phạm vi tiếp cận của các ứng viên. Họ có thể giảm chi phí đăng ký xuống 0 vì việc phân loại các hồ sơ dễ dàng đến mức không mất chi phí thực tế nào từ việc nhận nhiều hồ sơ hơn.

Cuối cùng, những sự thay đổi trong luồng công việc có thể trở nên quan trọng hơn. Với sự xếp hạng như vậy, trường có thể giảm thiểu thời gian giữa việc xem xét hồ sơ và gửi lời mời. Nếu sự xếp hạng đủ tốt, nó gần như có thể đồng thời thay đổi thời gian của toàn bộ luồng công việc và động lực cạnh tranh của những ứng viên MBA hàng đầu một cách đáng kể.

Loại AI này chỉ là giả định, nhưng ví dụ trên minh hoạ các công cụ AI khi được đặt trong một luồng công việc có thể khiến các công việc bị xoá bỏ (ví dụ, sự xếp hạng các hồ sơ một cách thủ công) cũng như thêm vào (ví dụ, sự quảng cáo tiếp cận rộng hơn). Mỗi doanh nghiệp sẽ có những kết quả khác nhau, nhưng khi chia nhỏ các luồng công việc, các doanh nghiệp có thể đánh giá liệu máy dự đoán có khả năng tiếp cận rộng hơn so với những quyết định cá nhân được thiết kế sẵn.

Các công cụ AI ảnh hưởng đến bàn phím iPhone như thế nào

Ở một mặt, bàn phím trên điện thoại thông minh của bạn có nhiều điểm chung với máy đánh chữ cơ khí hơn là bàn phím trên máy tính cá nhân. Nếu bạn đã sử dụng máy đánh chữ cơ khí, bạn sẽ biết rằng nếu bạn gõ quá nhanh, máy sẽ bị kẹt. Vì lí do này, bàn phím có bố cục QWERTY với tiêu chuẩn thiết kế hạn chế khả năng đánh hai phím liền kề, điều khiến những máy đánh chữ cơ khí mắc kẹt. Nhưng tính năng đó cũng làm giảm tốc độ của kể cả những người đánh máy nhanh nhất.

Thiết kế QWERTY đã tiếp tục tồn tại cho dù cơ chế gây ra những sự rắc rối đã không còn tồn tại nữa. Khi những kỹ sư Apple thiết kế iPhone, họ tranh luận xem liệu có thể loại bỏ hoàn toàn QWERTY không. Điều khiến họ quay lại sử dụng nó là do độ quen thuộc. Xét cho cùng, đối thủ cạnh tranh gần nhất lúc đó, BlackBerry, có một bàn phím QWERTY tốt đến mức sản phẩm

thường được gọi là “Crackberry” vì tính chất gây nghiện của nó.

“Dự án khoa học lớn nhất” của iPhone có lẽ là bàn phím mềm.[6] Nhưng vào cuối năm 2006 (iPhone được ra mắt vào năm 2007), bàn phím đó thực sự rất tệ. Nó không những không thể cạnh tranh với BlackBerry mà nó còn gây bực mình đến mức không ai có thể sử dụng nó để soạn thảo một tin nhắn văn bản, chứ chưa nói đến một email. Vấn đề là để phù hợp với màn hình LCD 4.7 inch, các phím phải rất nhỏ. Điều này có nghĩa là rất dễ gõ nhầm phím. Rất nhiều kỹ sư Apple đưa ra những thiết kế mà không liên quan đến QWERTY.

Chỉ có ba tuần để tìm ra một giải pháp – một giải pháp mà nếu không tìm ra, có thể sẽ giết chết toàn bộ dự án – mỗi nhà phát triển phần mềm của iPhone có quyền tự do khám phá những lựa chọn khác. Cho đến cuối tuần thứ ba, họ đã có một bàn phím nhìn giống như phiên bản thu nhỏ của QWERTY với một điều chỉnh đáng kể. Trong khi hình ảnh mà người dùng đã thấy không thay đổi, diện tích bề mặt xung quanh một tập hợp các phím được mở rộng khi đánh máy. Khi bạn gõ chữ “t”, khả năng cao là chữ cái tiếp theo sẽ là “h” và nên khu vực xung quanh phím đó sẽ được mở rộng. Tiếp theo đó, “e” và “I” được mở rộng, và tiếp tục như vậy.

Đây là kết quả của công cụ AI. Đi trước hầu hết tất cả mọi người, các kỹ sư ở Apple sử dụng máy tự học của năm 2006 để xây dựng những thuật toán dự đoán có thể khiến kích thước của phím thay đổi dựa vào điều mà một người gõ. Công nghệ với tính năng đó ảnh hưởng đến văn bản tự động chỉnh sửa mà bạn thấy ngày nay. Nhưng về cơ bản, lý do nó hoạt động là vì QWERTY. Bàn phím được thiết kế để đảm bảo bạn không phải gõ những phím liền kề sẽ cho phép các phím của điện thoại thông minh được mở rộng khi cần thiết bởi vì phím tiếp theo có thể sẽ không nằm gần phím bạn vừa sử dụng.

Điều mà các kỹ sư tại Apple đã làm khi phát triển iPhone là hiểu chính xác luồng công việc liên quan đến việc sử dụng bàn phím. Một người dùng phải xác định một phím, chạm vào nó và rồi rời đi tới một phím khác. Bằng việc chia nhỏ luồng công việc đó, họ nhận ra rằng một phím không cần phải giống nhau khi xác định và chạm vào. Quan trọng hơn, sự dự đoán có thể giải quyết làm thế nào để biết người dùng sẽ sử dụng phím gì tiếp. Hiểu được luồng công việc là quan trọng cho việc xác định cách tốt nhất để sử dụng công cụ AI. Điều này đúng với tất cả các luồng công việc.

NHỮNG ĐIỂM CHÍNH

• Các công cụ AI đều là các giải pháp mũi nhọn. Mỗi công cụ tạo ra một sự dự đoán nhất định và hầu hết được thiết kế để thực hiện một công việc nhất định. Nhiều công ty khởi nghiệp AI đã dự đoán việc xây dựng một công cụ AI riêng lẻ.

• Các tập đoàn lớn bao gồm các luồng công việc biến thông tin đầu vào thành thông tin đầu ra. Các luồng công việc được tạo thành bởi các nhiệm vụ (ví dụ, IPO của Goldman Sachs là một luồng công việc bao gồm 146 nhiệm vụ khác nhau). Trong việc quyết định cách bổ sung AI, các công ty sẽ chia nhỏ các luồng công việc thành các công việc, ước tính ROI cho việc xây dựng hoặc mua AI để thực hiện từng công việc, xếp hạng thứ tự AI liên quan đến ROI, sau đó bắt đầu từ đầu danh sách và đi xuống dưới. Đôi khi một công ty có thể mang một công cụ AI vào luồng công việc của họ và nhận ra lợi ích ngay lập tức do sự tăng năng suất của công việc đó. Tuy nhiên, thông thường nó không dễ dàng như vậy. Thu được lợi ích thực sự từ việc bổ sung một công cụ AI đòi hỏi nhiều sự suy nghĩ hoặc "tái cấu trúc" toàn bộ luồng công việc. Kết quả là, tương tự với cuộc cách mạng máy tính cá nhân, sẽ mất khá nhiều thời gian để thấy được năng suất đạt được từ AI trong nhiều công ty kinh doanh phổ biến.

• Để minh hoạ ảnh hưởng tiềm năng của AI trong một luồng công việc, chúng tôi miêu tả một AI viễn tưởng có thể dự đoán sự xếp hạng của bất kỳ hồ sơ MBA nào. Để thu được lợi ích từ máy dự đoán này, trường phải thiết kế lại luồng công việc của họ. Sẽ cần phải loại bỏ sự xếp hạng hồ sơ thủ công và tăng cường marketing chương trình, khi AI tăng lợi ích từ việc có nhiều ứng viên tốt hơn (sự dự đoán tốt hơn về việc ai sẽ thành công và giảm thiểu chi phí của việc đánh giá các hồ sơ). Trường sẽ phải chỉnh sửa việc đề xuất thêm các lời mời như học bổng và hỗ trợ tài chính vì sự gia tăng khả năng thành công của sinh viên. Cuối cùng, trường sẽ phải điều chỉnh những yếu tố khác của luồng công việc để tận dụng lợi thế của việc có khả năng cung cấp những quyết định nhập học ngay lập tức.

13Phân tích những quyết định

C

ác công cụ AI ngày nay khác xa với những máy có trí tuệ giống như con người của khoa học viễn tưởng (thường được gọi là “trí tuệ nhìn chung là nhân tạo” hoặc AGI, hoặc “AI mạnh”). Thế hệ AI hiện tại cung cấp những công cụ cho sự dự đoán là chủ yếu.

Quan điểm về AI này không làm suy giảm nó. Như Steve Jobs từng nhận xét, “Một trong những điều khiến chúng ta khác với những loài linh trưởng bậc cao là chúng ta là những người xây dựng công cụ.” Ông sử dụng ví dụ về chiếc xe đạp như là một công cụ cung cấp cho con người siêu năng lực trong việc di chuyển so với những loài vật khác. Và ông cảm thấy điều tương tự với máy tính: “Đối với tôi, máy tính là công cụ nổi bật nhất mà chúng ta từng sáng chế, và nó tương đương với một chiếc xe đạp trong suy nghĩ của chúng ta.”[1]

Các công cụ AI ngày nay dự đoán ý định của một văn bản (Echo của Amazon), dự đoán mệnh lệnh trong ngữ cảnh (Siri của Apple), dự đoán điều bạn muốn mua (gợi ý của Amazon), dự đoán những đường link kết nối bạn tới thông tin bạn muốn tìm (hệ thống tìm kiếm của Google), dự đoán khi nào cần sử dụng phanh để tránh tai nạn (Autopilot của Tesla) và dự đoán tin tức bạn muốn đọc (trang tin của Facebook). Tất cả những công cụ AI này đều không thực hiện toàn bộ luồng công việc. Thay vào đó, mỗi công cụ đóng vai trò là một thành phần dự đoán khiến công việc trở nên dễ dàng hơn cho ai cần đưa ra quyết định.

Nhưng làm thế nào để quyết định có nên dùng một công cụ AI nào đó cho một công việc nhất định trong doanh nghiệp của bạn? Mỗi công việc đều bao gồm nhiều sự dự đoán, và những quyết định này có một vài yếu tố dự đoán. Chúng tôi cung cấp cho bạn cách đánh giá AI trong bối cảnh của một công việc. Giống như khi chúng tôi gợi ý việc xác định công việc bằng cách chia nhỏ luồng công việc để xem AI có đóng vai trò gì không, bây giờ chúng tôi gợi ý phân tích những công việc thành những yếu tố thành phần.

Canvas AI

CDL đã tạo điều kiện cho chúng tôi tiếp xúc với nhiều công ty khởi nghiệp, tận dụng những công nghệ máy tự học gần đây để xây dựng các công cụ AI mới. Mỗi công ty trong phòng thí nghiệm đều dự đoán việc xây dựng một công cụ cụ thể, một số cho những trải nghiệm của khách hàng, nhưng đa số là cho những khách hàng doanh nghiệp. Loại thứ hai tập trung vào việc xác định những cơ hội công việc trong các luồng công việc để tập trung và xác định vị trí cung cấp của họ. Họ phân tích các luồng công việc, xác định một công việc với một yếu tố dự đoán, và xây dựng doanh nghiệp của họ dựa trên sự cung cấp công cụ để đưa ra sự dự đoán đó.

Khi tư vấn cho họ, chúng tôi cảm thấy hữu ích khi chia nhỏ một quyết định thành các yếu tố (liên quan đến hình 7-1): dự đoán, dữ liệu đầu vào, sự đào tạo, hành động, kết quả và phản hồi. Trong quá trình này, chúng tôi phát triển “canvas AI”, giúp phân tích các công việc để hiểu rõ vai trò tiềm năng của một máy dự đoán (xem hình 13-1).

Canvas này hỗ trợ cho việc dự tính, xây dựng và đánh giá các công cụ AI. Nó cung cấp một quy tắc để xác định mỗi thành phần trong quyết định của công việc. Nó yêu cầu sự mô tả mỗi thành phần rõ ràng.

Để xem cách nó hoạt động ra sao, hãy cùng xem xét công ty khởi nghiệp Atomwise, công ty cung cấp một công cụ dự đoán tập trung vào việc rút ngắn thời gian liên quan đến việc phát hiện những dược phẩm triển vọng. Hàng triệu những phân tử thuốc có thể trở thành thuốc, nhưng việc mua và thử nghiệm mỗi loại đều tốn thời gian cũng như tiền bạc. Vậy các công ty dược phẩm xác định loại thuốc cần thử nghiệm như thế nào? Họ sẽ đưa ra những dự đoán có cơ sở, dựa vào các nghiên cứu gợi ý những phân tử nào có khả năng cao trở thành thuốc tốt.

CEO của Atomwise Abraham Heifet đã cung cấp cho chúng tôi một giải thích nhanh về quá trình thử nghiện này rằng: “Để biết thuốc có hiệu quả không, nó phải kết dính với protein gây bệnh, và nó không được kết dính với protein trong gan, thận, tim, não và những bộ phận khác trong cơ thể mà có thể gây ra những tác dụng phụ độc hại. Quá trình này là về ‘kết dính những thứ bạn muốn nó kết dính, thất bại trong việc kết dính những thứ bạn không muốn.’”

Vậy nên, nếu các công ty dược phẩm có thể dự đoán khả năng kết dính, thì họ có thể xác định những phân tử nào có khả năng hiệu quả. Atomwise cung cấp sự dự đoán này bằng việc mang đến một công cụ AI khiến công việc xác định những loại thuốc tiềm năng trở nên hiệu quả hơn. Công cụ sử dụng AI để dự đoán khả năng kết dính của những phân tử, vậy nên Atomwise có thể gợi ý cho các công ty dược phẩm, trong danh sách xếp hạng, những phân tử nào có khả năng kết dính tốt nhất với một protein gây bệnh.

Ví dụ, Atomwise có thể cung cấp danh sách 20 phân tử có khả năng kết dính cao nhất, ví dụ cho virus Ebola. Thay vì kiểm tra từng phân tử một, máy dự đoán của Atomwise có thể xử lý hàng triệu khả năng. Mặc dù công ty dược phẩm vẫn cần phải thử nghiệm và xác nhận những loại phân tử tiềm năng thông qua sự đánh giá và hoạt động của con người cũng như máy móc, công cụ AI của Atomwise giảm thiểu chi phí và tăng tốc độ tìm ra những loại phân tử phù hợp.

Vậy sự đánh giá này xuất hiện ở đâu? Khi nhận ra giá trị tổng hợp của một phân tử tiềm năng cụ thể đối với ngành công nghiệp dược phẩm. Giá trị này có hai dạng: xác định bệnh và hiểu rõ những tác dụng phụ tiềm năng. Trong việc lựa chọn các phân tử để thử nghiệm, công ty cần xác định những sự trả giá của việc xác định bệnh và chi phí cho những tác dụng phụ. Như Heifets nhận xét, “Bạn chấp nhận những tác dụng phụ của hoá trị hơn là kem trị mụn.”

Máy dự đoán của Atomwise học từ dữ liệu về khả năng kết dính. Tính đến tháng 7 năm 2017, có đến 38 triệu điểm dữ liệu công khai về khả năng kết dính cộng thêm nhiều thông tin nó mua hoặc tự học. Mỗi điểm dữ liệu bao gồm những đặc tính của phân tử và protein cũng như sự đánh giá khả năng kết dính giữa các phân tử và protein. Khi Atomwise đưa ra nhiều đề xuất hơn, nó có thể nhận thêm nhiều phản hồi từ khách hàng, vậy nên máy dự đoán sẽ tiếp tục được cải thiện.

Bằng cách sử dụng máy này, với dữ liệu về các đặc tính của protein, Atomwise có thể dự đoán những phần tử nào có khả năng kết dính cao nhất. Nó cũng có thể lấy dữ liệu về những đặc tính của protein và dự đoán những phần tử nào chưa từng được sản xuất nhưng có khả năng kết dính cao nhất.

Cách để phân tích việc lựa chọn phần tử của Atomwise là điền vào canvas

(xem hình 13-2). Điều này đồng nghĩa với việc xác định những điều sau:

• Hành động: Bạn đang cố làm điều gì? Với Atomwise là để thử nghiệm các phần tử giúp chữa trị hoặc phòng ngừa bệnh.

• Dự đoán: Bạn cần biết điều gì để đưa ra quyết định? Atomwise dự đoán khả năng kết dính của những phân tử và protein tiềm năng.

• Đánh giá: Làm thế nào để đánh giá các kết quả và sai sót khác nhau? Atomwise và khách hàng của họ đặt tiêu chí này nhắm đến sự quan trọng của việc xác định bệnh và những chi phí liên quan của những tác dụng phụ tiềm năng.

• Kết quả: Tiêu chuẩn của một công việc thành công là gì? Với Atomwise, đó là kết quả của việc thử nghiệm. Cuối cùng, liệu việc thử nghiệm có dẫn đến việc ra đời của một loại thuốc mới?

• Thông tin đầu vào: Bạn cần dữ liệu nào để chạy thử thuật toán dự đoán? Atomwise sử dụng dữ liệu về những đặc tính của protein bệnh để dự đoán.

• Đào tạo: Bạn cần loại dữ liệu nào để đào tạo thuật toán dự đoán? Atomwise sử dụng dữ liệu về khả năng kết dính cùng những đặc tính của phân tử và protein.

• Phản hồi: Làm thế nào để bạn có thể sử dụng những kết quả để cải thiện thuật toán? Atomwise sử dụng các kết quả thử nghiệm, cho dù là sự thành công, để cải thiện những dự đoán trong tương lai.

Đề xuất giá trị của Atomwise nằm ở việc cung cấp một công cụ AI hỗ trợ công việc dự đoán trong luồng công việc khám phá thuốc cho khách hàng của họ. Nó loại bỏ công việc dự đoán khỏi tầm tay của con người. Để cung cấp giá trị như vậy, nó đã tích luỹ một lượng dữ liệu đặc biệt để dự đoán khả năng kết dính. Giá trị của sự dự đoán nằm ở việc giảm thiểu chi phí và tăng khả năng thành công cho sự phát triển thuốc. Những khách hàng của Atomwise sử dụng sự dự đoán kết hợp với đánh giá chuyên sâu của họ về những sự trả giá cho các phân tử với khả năng kết dính khác nhau với nhiều loại protein khác nhau.

Canvas AI cho việc tuyển sinh MBA

Canvas cũng hữu ích trong những tổ chức lớn. Để áp dụng nó, chúng tôi chia nhỏ luồng công việc thành những mục tiêu. Ở đây, chúng tôi xem xét canvas AI tập trung vào việc nên lựa chọn những ứng viên MBA phù cho chương trình. Hình 13-3 cung cấp một canvas khả thi.



Vậy canvas này tới từ đâu? Đầu tiên, việc tuyển sinh đòi hỏi sự dự đoán: Ai sẽ là sinh viên tốt nhất hoặc có giá trị cao? Điều đó dường như rất đơn giản. Chúng ta đơn giản chỉ cần định nghĩa từ “tốt nhất”. Chiến lược của trường có thể giúp xác định được điều này. Tuy nhiên, nhiều tổ chức có những tuyên bố nhiệm vụ mơ hồ, khiến họ xuất hiện tốt hơn trên những tờ quảng cáo nhưng lại không tốt khi xác định mục tiêu dự đoán cho AI.

Các trường kinh tế có nhiều chiến lược không rõ ràng hoặc rõ ràng định nghĩa “sự tốt nhất” của họ. Chúng có thể là những chỉ dẫn đơn giản, ví dụ tối đa hóa các điểm bài thi chuẩn hóa như GMAT hay những mục tiêu lớn hơn như tuyển những sinh viên có thể sẽ giúp trường tăng thứ hạng trên tờ *Financial Times* hay *US News & World Report*. Họ có thể cũng muốn sinh viên của họ có nhiều kỹ năng định tính và định lượng. Hoặc họ có thể muốn nhiều sinh viên quốc tế. Hoặc họ có thể muốn sự đa dạng. Không trường nào có thể theo đuổi tất cả những mục tiêu này cùng lúc và thường phải đưa ra lựa chọn. Nếu không họ sẽ thỏa hiệp trên tất cả các khía cạnh nhưng không nổi trội ở khía cạnh nào.

Trong hình 13-3, chúng tôi hình dung ra chiến lược của trường là có được sức ảnh hưởng lớn nhất lên kinh tế toàn cầu. Ý kiến chủ quan này là chiến lược bởi nó mang tính toàn cầu thay vì cục bộ và tìm kiếm sự ảnh hưởng thay vì tối đa hóa tài chính của sinh viên hoặc trở nên giàu có.

Để AI dự đoán được ảnh hưởng kinh tế toàn cầu, chúng ta cần đo lường nó. Ở đây, chúng tôi giả sử đóng vai trò của RFE. Chúng ta có loại dữ liệu đào tạo nào mà có thể đại diện cho ảnh hưởng kinh tế toàn cầu? Một lựa chọn có thể là xác định cựu sinh viên giỏi nhất từ mỗi khóa – 50 cựu sinh viên mỗi năm mà có ảnh hưởng lớn nhất. Việc lựa chọn những cựu sinh viên, đương nhiên, là chủ quan, nhưng không phải là không thể.

Trong khi chúng ta đặt ảnh hưởng kinh tế toàn cầu là mục tiêu cho máy dự đoán, giá trị của việc chấp nhận một sinh viên cụ thể là vấn đề của sự đánh giá. Các trường sẽ phải trả giá ra sao nếu dự đoán sai và nhận một sinh viên yếu kém trong môi trường sinh viên danh giá? Các trường sẽ trả giá ra sao nếu từ chối một sinh viên bị dự đoán nhầm là yếu kém? Việc ước lượng sự đánh đổi đó là "sự đánh giá", một yếu tố rõ ràng trong canvas AI. Một khi chúng ta xác định được mục tiêu của sự dự đoán, việc xác định dữ liệu đầu vào sẽ trở nên đơn giản. Chúng ta cần thông tin hồ sơ của những tân sinh viên để dự đoán xem họ sẽ như thế nào. Chúng ta có thể sử dụng cả truyền thông xã hội. Theo thời gian, chúng ta sẽ quan sát sự nghiệp của các sinh viên và có thể sử dụng phản hồi đó để cải thiện những sự dự đoán. Những sự đự đoán sẽ nói cho chúng ta biết những ứng viên nào nên được chấp nhận, nhưng chỉ sau khi xác định được mục tiêu của chúng ta và đánh giá chi phí của việc gây ra một sai sót.

NHỮNG ĐIỂM CHÍNH

• Các công việc cần được phân tích để thấy ở giai đoạn nào thì nên sử dụng máy dự đoán. Điều này cho phép bạn ước tính lợi ích của việc cải thiện sự dự đoán và chi phí của việc tạo ra loại dự đoán đó. Một khi đã tạo ra những ước tính hợp lý, hệ thống sẽ xếp hạng thứ tự của AI từ ROI cao nhất đến thấp nhất, bắt đầu từ trên đầu và đi xuống dưới, bổ sung các công cụ AI chừng nào ROI theo dự đoán vẫn hợp lý.

• Canvas AI hỗ trợ quá trình phân tích. Điền vào canvas AI cho mỗi quyết định hoặc công việc. Việc này sẽ đưa quy tắc và cấu trúc đi vào quy trình. Nó bắt bạn phải nắm rõ về ba loại dữ liệu được yêu cầu: đào tạo, đầu vào và phản hồi. Nó cũng bắt bạn phải nói chính xác bạn cần dự đoán điều gì, sự đánh giá cần thiết để ước lượng giá trị liên quan đến những hành động và kết quả khác nhau, những khả năng hành động và những kết quả dự kiến.

• Cốt lõi của canvas AI là sự dự đoán. Bạn cần xác định được dự đoán cốt lõi của công việc, và điều này có thể đòi hỏi sự thấu hiểu sâu sắc về AI. Nỗ lực để trả lời câu hỏi này thường được bắt đầu với một cuộc thảo luận giữa đội ngũ lãnh đạo: "Mục tiêu thực tế của chúng ta là gì?" Sự dự đoán đòi hỏi một sự xác định cụ thể mà thường không thấy trên những tuyên bố mục đích. Ví dụ với một trường kinh tế, có thể dễ dàng nói là họ tập trung vào việc tuyển những sinh viên "tốt nhất", nhưng để xác định cụ thể sự dự đoán, chúng ta

cần xác định “sự tốt nhất” là gì – lời mời với mức lương cao nhất sau khi tốt nghiệp? Khả năng cao đảm nhận vai trò CEO trong vòng năm năm? Đa tài nhất? Có khả năng đóng góp cho trường sau khi tốt nghiệp?

Ngay cả những mục tiêu dường như là đơn giản nhất, cũng không đơn giản như chúng ta nghĩ. Chúng ta có nên dự đoán hành động cần làm để tối đa hóa lợi nhuận trong tuần đó, quý đó, năm đó, hay thập kỷ đó không? Các công ty thường phải quay lại những điều cơ bản nhất để điều chỉnh những mục tiêu của họ và cụ thể hóa tuyên bố mục tiêu của họ như là bước đầu tiên trong việc thực hiện chiến lược AI.

14Thiết kế lại công việc

T

rước khi AI ra đời và mạng Internet là cuộc cách mạng máy tính, máy tính tạo ra thuật toán – đặc biệt, thêm nhiều điều khác nữa – có giá thành rẻ. Một trong những ứng dụng hàng đầu đã giúp việc bảo quản sách dễ dàng hơn.

Kỹ sư máy tính Dan Bricklin đã có ý tưởng này khi là một sinh viên MBA, ông cảm thấy bực bội khi phải thực hiện những tính toán lặp đi lặp lại để đánh giá những viễn cảnh khác nhau của Trường Kinh tế Harvard. Vậy nên, ông đã viết một chương trình máy tính để thực hiện những tính toán đó và nhận ra nó hữu ích đến mức ông cùng với Bob Frankston đã phát triển nó thành VisiCalc cho máy tính Apple II. VisiCalc là ứng dụng hàng đầu trong kỷ nguyên máy tính cá nhân và là lý do khiến nhiều doanh nghiệp lần đầu tiên mua máy tính cho những văn phòng của họ.[1] Ứng dụng này không những giảm thiểu thời gian tính toán đến 100 lần, mà còn cho phép các doanh nghiệp phân tích nhiều viễn cảnh hơn.

Ở thời điểm đó, những người được giao nhiệm vụ tính toán hoạt động là những chuyên viên lưu trữ; vào cuối những năm 1970, có hơn 400.000 người làm việc ở Hoa Kỳ. Bảng tính đã loại bỏ việc khiến họ tốn thời gian nhất – thuật toán. Bạn có thể nghĩ rằng những chuyên viên lưu trữ sẽ bị thất nghiệp. Nhưng chúng tôi chưa nghe thấy bài ca than phiền mất việc của những chuyên viên lưu trữ và không có sự phản đối nào tạo ra rào cản cho việc sử dụng rộng rãi của bảng tính. Vậy tại sao những chuyên viên lưu trữ không coi bảng tính là một mối đe dọa?

Bởi vì VisiCalc khiến họ trở nên có ích hơn. Nó khiến sự tính toán trở nên dễ dàng. Bạn có thể dễ dàng đánh giá được lợi nhuận dự kiến và cách nó thay đổi nếu bạn thay đổi nhiều giả định. Thay vì dự đoán liệu sự đầu tư này có hiệu quả hay không, bạn có thể so sánh với những đầu tư khác nhau qua các dự đoán khác nhau và lựa chọn sự đầu tư tốt nhất. Một bảng tính có thể cho bạn những câu trả lời một cách dễ dàng, và làm tăng lợi nhuận. Những người đã từng tính toán câu trả lời trước khi bảng tính ra đời là những người phù hợp nhất để hỏi về bảng tính đã được máy tính hóa. Họ không bị thay thế mà

còn được trợ giúp bởi sức mạnh siêu phàm.

Loại viễn cảnh này – một công việc được trợ giúp khi máy móc bắt đầu thực hiện một số, chứ không phải tất cả, các công việc – có khả năng trở nên khá phổ biến như là một hệ quả tự nhiên của việc bổ sung các công cụ AI. Các nhiệm vụ trong công việc sẽ thay đổi. Một số sẽ bị xóa bỏ khi máy dự đoán thực hiện chúng, một số sẽ được thêm vào khi con người có nhiều thời gian hơn. Và với nhiều công việc, những kỹ năng cần thiết trước đây sẽ thay đổi và những kỹ năng mới sẽ thay thế. Giống như những chuyên viên lưu trữ trở nên thành thạo bảng tính, sự thiết kế lại công việc bởi các công cụ AI cũng sẽ rất mạnh mẽ.

Quá trình bổ sung các công cụ AI sẽ quyết định các kết quả mà bạn muốn nhấn mạnh. Nó bao gồm việc đánh giá toàn bộ các luồng công việc, liệu chúng ở trong hay phân bố rải rác trong các công việc (các bộ phận hoặc tổ chức), sau đó phân chia luồng công việc thành những công việc nhỏ và xem liệu bạn có thể sử dụng máy dự đoán hiệu quả trong các công việc này hay không. Cuối cùng, bạn phải sắp xếp lại những phần việc đó thành những công việc hoàn chỉnh.

Thiếu sự liên kết trong sự tự động hóa

Trong một vài trường hợp, mục tiêu là tự động hóa toàn bộ các nhiệm vụ liên quan đến công việc. Các công cụ AI không có khả năng tự hoàn thành điều này vì các luồng công việc chịu trách nhiệm cho sự tự động hóa có hàng loạt các nhiệm vụ không thể (dễ dàng) tránh khỏi, ngay cả với những nhiệm vụ tưởng chừng không yêu cầu kỹ năng cao và không quan trọng.

Để tự động hóa hoàn toàn một nhiệm vụ, một phần không hoạt động có thể làm hỏng toàn quá trình. Bạn cần xem xét từng bước một. Những nhiệm vụ nhỏ đó có thể là những liên kết còn thiếu trong sự tự động hóa và có thể hạn chế cách cải thiện công việc. Do đó, các công cụ AI giải quyết những liên kết còn thiếu này có thể có những ảnh hưởng quan trọng.

Hãy xem xét ngành công nghiệp hậu cần, ngành đã phát triển nhanh chóng trong suốt hai thập kỷ qua nhờ sự phát triển nhanh chóng của mua sắm trực tuyến. Hậu cần là một bước quan trọng trong việc bán lẻ nói chung và trong thương mại điện tử nói riêng. Đây là quá trình lấy một đơn hàng và xử lý nó

bằng cách vận chuyển hàng đến với khách hàng. Trong thương mại điện tử, hậu cần bao gồm những bước như xác định sản phẩm trong một cơ sở nhà kho lớn, lựa chọn những sản phẩm trên kệ, quét chúng để quản lý hàng tồn kho, đặt chúng vào giỏ hàng, đóng gói chúng trong hộp, dán nhãn cho hộp và vận chuyển chúng đi. Nhiều ứng dụng ban đầu của máy tự học cho giai đoạn hậu cần liên quan đến quản lý hàng tồn kho: dự đoán những sản phẩm nào sẽ bán hết, những sản phẩm không cần đặt hàng lại vì nhu cầu mua thấp... Những dự đoán này đã trở thành một phần quan trọng của bán lẻ trực tiếp và quản lý kho hàng trong nhiều thập kỷ. Những công nghệ máy tự học đã khiến những dự đoán này trở nên tốt hơn.

Trong suốt hai thập kỷ qua, đa số những quá trình hậu cần còn lại đã được tự động hóa. Ví dụ, nghiên cứu đã xác định rằng các công nhân ở trung tâm hậu cần dành hơn nửa thời gian của họ đi khắp kho hàng để tìm những sản phẩm và đặt chúng vào giỏ hàng. Kết quả là nhiều công ty phát triển một quy trình tự động hóa để mang những kệ hàng đến chỗ các công nhân, từ đó giảm thiểu thời gian họ phải đi lại. Amazon đã mua lại công ty hàng đầu trong thị trường này, Kiva, vào năm 2012 với 775 triệu đô la và cuối cùng ngừng cung cấp dịch vụ cho những khách hàng Kiva khác. Những nhà cung cấp khác lần lượt nổi lên để đáp ứng nhu cầu của các trung tâm hậu cần trong nhà và các công ty hậu cần thứ ba.

Mặc dù có sự tự động hóa đáng kể, các trung tâm hậu cần vẫn cần tuyển nhiều nhân công. Về cơ bản, trong khi các robot có thể lấy một vật và chuyển nó tới tay con người, một ai đó vẫn cần phải "nhận đồ" – tức là xác định xem vật gì sẽ tới chỗ nào, rồi nhấc vật đó lên và di chuyển đi. Mảnh ghép cuối cùng là thách thức lớn nhất bởi vì nó khá khó nắm bắt. Miễn là con người vẫn đóng vai trò này, các nhà kho không thể tận dụng hoàn toàn lợi thế của sự tự động hóa bởi vì họ vẫn cần con người, với không gian đi lại, phòng nghỉ, phòng vệ sinh, giám sát chống trộm… Tất cả những điều đó đều tốn kém.

Vai trò của con người vẫn được tiếp tục trong việc hậu cần đặt hàng là do hiệu suất liên quan đến việc nắm bắt, nhận vật gì đó và đặt nó ở một nơi khác. Công việc này cho đến nay đã vượt qua sự tự động hóa. Kết quả là Amazon có tới 4.000 nhân công nhận đồ toàn thời gian và hàng chục nghìn người khác làm việc bán thời gian trong mùa nghỉ bận rộn. Một người nhận đồ trung bình xử lý 120 đơn hàng mỗi giờ. Nhiều công ty xử lý khối lượng

hậu cần lớn sẽ muốn tự động hóa sự nhận đồ. Trong ba năm qua, Amazon đã khuyến khích đội ngũ robot tốt nhất trên thế giới nghiên cứu vấn đề này bằng việc tổ chức Amazon Picking Challenge, tập trung vào việc tự động hóa nhận hàng trong những môi trường nhà kho không có cấu trúc chuẩn. Mặc dù những đội ngũ hàng đầu đến từ những tổ chức như MIT đã nghiên cứu vấn đề này, nhiều đội đã sử dụng thiết bị công nghiệp tân tiến từ Baxter, Yaskawa Motoman, Universal Robots, ABB, PR2 và Barrett Arm, nhưng cho đến thời điểm viết cuốn sách này, họ vẫn chưa giải quyết được vấn đề này một cách thỏa đáng.

Robot hoàn toàn có khả năng lắp ráp một chiếc xe hoặc lái một chiếc máy bay. Vậy sao chúng không thể nhận đồ từ nhà kho của Amazon và xếp đồ vào hộp? Công việc này nghe thật đơn giản khi đem ra so sánh. Robot có thể lắp ráp một máy tự động vì những thành phần đã được chuẩn hóa cao và quy trình được lập trình cao. Tuy nhiên, một nhà kho của Amazon đa dạng về kích thước, hình dạng, trọng lượng, số lượng sản phẩm đặt trên giá ở nhiều vị trí khác nhau và định hướng cho những vật thể không phải hình chữ nhật. Nói cách khác, vấn đề nhận đồ trong một nhà kho được đặc trưng hóa bởi vô số “nếu”, trong khi việc lắp ráp xe được thiết kế với rất ít “nếu”. Vậy nên, để có thể nhận đồ trong không gian nhà kho, robot cần phải có khả năng “nhìn” vật (phân tích hình ảnh), dự đoán góc độ và áp suất thích hợp để giữ được vật mà không rơi hay làm vỡ nó. Nói cách khác, dự đoán là vấn đề gốc rễ của việc nhận nhiều loại đồ trong trung tâm hậu cần.

Các nghiên cứu về vấn đề nhận đồ sử dụng phương pháp học tăng cường để đào tạo robot bắt chước con người. Công ty khởi nghiệp ở Vancouver – Kindred – được thành lập bởi Suzanne Gildert, Geordie Rose và một nhóm bao gồm cả một trong số các tác giả (Ajay) – đang sử dụng một chú robot tên là Kindred Sort, cánh tay có sự kết hợp giữa phần mềm tự động hóa và bộ điều khiển do con người.[2] Sự tự động hóa xác định đối tượng và nơi mà nó muốn đến, trong khi con người – đeo một bộ thực tế ảo – hướng dẫn cánh tay robot nhận đồ và di chuyển chúng.

Trong quá trình lặp lại đầu tiên, con người có thể ngồi ở bất kỳ đâu xa nhà kho và hoàn thiện liên kết còn thiếu trong luồng công việc hậu cần, quyết định góc độ tiếp cận và áp lực bám, thông qua sự vận hành của cánh tay robot. Tuy nhiên về lâu dài, Kindred đang sử dụng một máy dự đoán được

đào tạo bằng cách quan sát cách con người nhận vật thông qua bộ điều khiển từ xa để dạy robot tự làm việc đó.

Chúng ta có nên ngừng việc đào tạo các bác sĩ X-quang?

Vào tháng 10 năm 2016, khi đứng trên sân khấu trước 600 khán giả ở buổi hội thảo hằng năm CDL về vấn đề kinh doanh trí tuệ nhân tạo, Geoffrey Hinton – một người tiên phong trong hệ thống mạng lưới học sâu – tuyên bố: "Chúng ta nên ngừng đào tạo các bác sĩ X-quang ngay bây giờ." Theo quan điểm của Hinton, AI sẽ sớm có thể xác định những điểm bất thường trong hình ảnh y khoa giỏi hơn bất kỳ người nào. Các bác sĩ X-quang đã lo sợ máy móc có thể thay thế họ kể từ đầu những năm 1960.[3] Vậy điều gì khiến công nghệ ngày nay trở nên khác biệt? Các kỹ thuật máy tự học đều đang giỏi hơn trong việc dự đoán thông tin còn thiếu, bao gồm sự nhận dạng và nhận diện các vật trong hình ảnh. Với một bộ hình ảnh mới, các kỹ thuật viên có thể so sánh hiệu quả hàng triệu ví dụ trong quá khứ một cách dễ dàng hoặc không và dự đoán liệu hình ảnh mới có gợi ý sự xuất hiện của bệnh không. Kiểu nhận dạng mẫu để dự đoán bệnh này là những gì mà các bác sĩ X-quang làm.[4]

IBM với hệ thống Watson, và nhiều công ty khởi nghiệp khác đã thương mại hóa các công cụ AI vào trong X-quang. Watson có thể xác định một sự thuyên tắc động mạch phổi và những vấn đề liên quan đến tim khác. Một công ty khởi nghiệp khác, Entitic, sử dụng việc học sâu để phát hiện các nốt phổi (một việc được thực hiện khá thường xuyên) và cả vết rạn xương (phức tạp hơn). Những công cụ mới này là cốt lõi cho sự dự đoán của Hinton nhưng là chủ đề thảo luận giữa những bác sĩ X-quang và những nhà bệnh lý học.[5]

Vậy cách tiếp cận của chúng tôi tiết lộ điều gì về tương lai của các bác sĩ X-quang? Các bác sĩ X-quang sẽ dành ít thời gian đọc hình ảnh hơn. Dựa vào những cuộc phỏng vấn với những bác sĩ chăm sóc và các bác sĩ X-quang, cũng như dựa vào kiến thức về những nguyên tắc kinh tế được xây dựng chắc chắn, chúng tôi miêu tả những vai trò quan trọng chỉ dành cho các chuyên gia trong lĩnh vực hình ảnh y khoa.[6]

Đầu tiên, và có lẽ là hiển nhiên nhất, con người vẫn cần phải xác định hình ảnh cho một bệnh nhân nào đó. Hình ảnh y khoa là việc rất tốn kém, cả về

thời gian và những hệ quả sức khỏe tiềm ẩn của việc tiếp xúc với bức xạ. Khi chi phí của việc phân tích hình ảnh giảm, số lượng hình ảnh được phân tích sẽ gia tăng, vậy nên có khả năng rằng trong thời gian ngắn và tương đối trung bình, sự gia tăng này sẽ đánh dấu sự giảm thiểu thời gian của con người dành cho mỗi hình ảnh.

Thứ hai, có nhiều bác sĩ chẩn đoán X-quang và bác sĩ can thiệp X-quang. Những sự tiến bộ trong việc nhận dạng đối tượng mà có thể thay đổi bản chất của X-quang đều liên quan đến sự chẩn đoán X-quang. Sự can thiệp X-quang sử dụng những hình ảnh thời gian thực để hỗ trợ những thủ tục y tế. Hiện nay, điều này bao gồm sự đánh giá của con người và hành động khéo léo của con người mà không bị ảnh hưởng bởi những sự tiến bộ của AI, ngoại trừ có lẽ nó đã khiến công việc của bác sĩ can thiệp X-quang dễ dàng hơn bằng cách cung cấp những hình ảnh dễ xác định hơn.

Thứ ba, nhiều bác sĩ X-quang tự thấy bản thân họ là “bác sĩ của bác sĩ”.[7] Một phần quan trọng trong công việc của họ là truyền đạt được ý nghĩa của những hình ảnh tới những bác sĩ chăm sóc. Phần thách thức là việc diễn giải các hình ảnh X-quang (“nghiên cứu”, theo ngôn ngữ của họ) thường mang tính xác suất: “Có 70% khả năng đây là bệnh X, có 20% không có bệnh và 10% khả năng là bệnh Y. Tuy nhiên, nếu hai tuần kể từ bây giờ, triệu chứng này lại xuất hiện, thì có đến 99% khả năng là bệnh X và 1% là không có bệnh.” Nhiều bác sĩ chăm sóc không giỏi trong việc thống kê số liệu và gặp khó khăn trong việc diễn giải các khả năng và các khả năng có điều kiện. Các bác sĩ X-quang giúp họ diễn giải những con số để các bác sĩ chăm sóc có thể làm việc với bệnh nhân và đưa ra quyết định hành động tốt nhất. Qua thời gian, AI sẽ cung cấp các khả năng, nhưng ít nhất là trong thời gian ngắn và có khả năng trung bình, bác sĩ X-quang vẫn sẽ đóng vai trò trong việc diễn giải thông tin đầu ra của AI cho bác sĩ chăm sóc.

Thứ tư, các bác sĩ X-quang sẽ giúp đào tạo máy móc diễn giải các hình ảnh từ những thiết bị hình ảnh mới khi công nghệ cải thiện. Một vài chuyên gia cũng là bác sĩ X-quang cực kỳ nổi tiếng, những người diễn giải hình ảnh và giúp máy học cách chẩn đoán, sẽ có vai trò này. Thông qua AI, các bác sĩ X-quang sẽ tận dụng những kỹ năng tuyệt vời của họ trong việc chẩn đoán để đào tạo máy. Những dịch vụ của họ sẽ có giá trị cao. Thay vì được trả tiền bởi các bệnh nhân, họ có thể được đền bù bởi các kỹ thuật họ mới dạy cho AI hoặc

cho mỗi bệnh nhân được thử nghiệm trên AI họ đào tạo.[8]

Máy dự đoán sẽ giảm sự không chắc chắn, nhưng nó sẽ không hoàn toàn loại bỏ nó. Ví dụ, máy có thể đưa ra một sự dự đoán như dưới đây:

Dựa trên các thông tin nhân khẩu học và phân tích hình ảnh của ông Patel, khối u có 66.6% khả năng là u lành tính, 33.3% khả năng là u ác tính, và 0.1% là không có u.

Nếu như máy dự đoán đưa ra một dự đoán đơn giản – lành tính hay không – mà không có chỗ cho sai sót, việc chúng ta cần phải làm gì sẽ rất đơn giản. Ở thời điểm này, bác sĩ cần xem xét có nên thực hiện một thủ tục can thiệp không, ví dụ như sinh thiết, để tìm hiểu thêm. Việc thực hiện sinh thiết là một quyết định ít rủi ro hơn; nó tốn kém, nhưng nó có thể cho ra một chẩn đoán chắc chắn hơn.

Chúng ta có thể nhìn thấy được rằng vai trò của máy dự đoán là tăng sự tự tin của bác sĩ trong việc không phải thực hiện sinh thiết. Những thủ tục như vậy sẽ ít tốn kém hơn (đặc biệt là với các bệnh nhân). Chúng thông báo cho các bác sĩ liệu bệnh nhân có thể tránh được kiểm tra mang tính đòi hỏi (như sinh thiết) hay không và giúp họ tự tin trong việc không cần phương pháp điều trị và phân tích chuyên sâu hơn. Nếu như máy cải thiện sự dự đoán, nó sẽ dẫn đến ít xét nghiệm tốn kém hơn.

Vì vậy, vai trò thứ năm và cũng là cuối cùng của những chuyên gia trong lĩnh vực phân tích hình ảnh y khoa là sự đánh giá trong việc quyết định xem có cần thực hiện xét nghiệm tốn kém hay không, ngay cả khi máy gợi ý khả năng cao rằng không có vấn đề gì. Bác sĩ có thể có thông tin về sức khỏe tổng quan của bệnh nhân, sự căng thẳng thần kinh do sự bi quan sai lầm, hoặc một số dữ liệu định tính khác. Những thông tin như vậy có thể không dễ mã hóa và có sẵn cho máy, đồng thời có thể đòi hỏi một cuộc trò chuyện giữa bác sĩ X-quang với kiến thức chuyên môn trong việc diễn giải những khả năng và bác sĩ chăm sóc biết rõ nhu cầu của bệnh nhân. Thông tin này có thể dẫn đến việc con người không cần sự gợi ý của AI để thực hiện hành động.

Do vậy, năm vai trò rõ ràng của con người trong việc sử dụng hình ảnh y khoa sẽ tiếp tục, ít nhất là trong ngắn hạn và trung hạn: lựa chọn hình ảnh, sử dụng những hình ảnh thời gian thực trong những thủ tục y khoa, diễn giải

thông tin đầu ra của máy, đào tạo máy sử dụng những công nghệ mới và tận dụng sự đánh giá không cần gợi ý của máy dự đoán mà dựa vào những thông tin không có sẵn với máy. Việc liệu các bác sĩ X-quang có tương lai hay không phụ thuộc vào việc họ có ở vị trí tốt nhất để thực hiện những vai trò đó hay không, nếu những chuyên gia khác thay thế họ, hoặc nếu những lớp công việc mới phát triển, như sự kết hợp giữa bác sĩ X-quang/bác sĩ bệnh lý học (ví dụ, một bác sĩ X-quang nhưng cũng phân tích sinh thiết, ngay sau khi phân tích hình ảnh).[9]

Hơn cả một người lái xe

Một vài công việc có thể tiếp tục tồn tại nhưng sẽ đòi hỏi những kỹ năng mới. Việc tự động hóa một công việc cụ thể có thể nhấn mạnh những nhiệm vụ quan trọng với một công việc nhưng trước đây bị đánh giá thấp. Hãy xem xét trường hợp của người lái xe buýt trường học. Phần "lái xe" trong công việc bao gồm lái một chiếc xe buýt từ nhà đến trường và ngược lại. Với sự ra đời của những xe tự lái và sự tự động hóa lái xe, công việc của người lái xe buýt có thể sẽ biến mất. Khi các giáo sư của Đại học Oxford, Carl Frey và Michael Osborne xem xét những loại kỹ năng cần để thực hiện một công việc, họ kết luận rằng những người lái xe buýt trường học (để phân biệt với những người lái xe buýt công cộng bình thường) có 89% khả năng bị thay thế bởi sự tự động hóa trong vòng một hoặc hai thập kỷ tới.[10]

Khi một người "người lái xe buýt trường học" không còn lái xe buýt đến trường nữa, phải chăng chính quyền địa phương nên bắt đầu tiết kiệm khoản lương này? Ngay cả khi một chiếc xe buýt có thể tự lái, những người lái xe buýt có thể làm được nhiều hơn việc lái xe. Đầu tiên, họ là những người lớn, chịu trách nhiệm giám sát một nhóm nhiều học sinh để bảo vệ chúng khỏi những nguy hiểm bên ngoài xe buýt. Thứ hai, quan trọng không kém, họ có trách nhiệm giữ kỷ luật trong xe buýt. Sự đánh giá của con người trong việc quản lý trẻ em và bảo vệ chúng vẫn cần thiết. Việc xe có thể tự lái không loại bỏ những công việc bên lề này, nhưng nó đồng nghĩa với việc người lớn trên xe buýt có thể tập trung vào những công việc đó hơn.

Vậy nên kỹ năng của "những người lái xe buýt trường học" sẽ thay đổi. Những người lái xe sẽ đóng vai trò giống như giáo viên hơn. Nhưng điểm trọng yếu ở đây là sự tự động hóa loại bỏ con người khỏi một nhiệm vụ

nhưng không nhất thiết loại bỏ họ khỏi một công việc. Sự tự động hóa những công việc khiến chúng ta suy nghĩ cẩn thận hơn về những điều tạo nên một công việc, điều mà con người đang thực sự làm. Giống như những người lái xe buýt trường học, những người lái xe tải dài làm nhiều hơn là mỗi việc lái xe. Lái xe tải là một trong những danh mục phân loại công việc lớn nhất ở Hoa Kỳ và thường xuyên là một ứng viên tiềm năng cho sự tự động hóa

Nhưng liệu chúng ta có thực sự thấy những chiếc xe tải di chuyển trên khắp lục địa mà không cần con người kiểm soát? Hãy nghĩ về những thách thức được đặt ra vì đa số những xe tải này sẽ khó có thể xa rời sự giám sát của con người. Ví dụ, những chiếc xe tải với những món hàng sẽ có khả năng bị cướp trộm. Những chiếc xe tải như vậy có thể sẽ không hoạt động được nếu không có con người và dễ trở thành một mục tiêu.

Giải pháp rất rõ ràng: một người sẽ đi cùng theo xe. Công việc đó sẽ dễ dàng hơn nhiều so với việc lái xe, đồng thời cho phép xe lái trong thời gian dài hơn mà không cần nghỉ ngơi. Con người có thể di chuyển với một chiếc xe lớn hơn hoặc thậm chí là một đoàn xe.[11] Nhưng ít nhất một chiếc xe tải trong đoàn vẫn sẽ có một khoang ngồi cho con người để bảo vệ xe, xử lý những tình huống hậu cần cũng như những việc liên quan đến chất hàng và dỡ hàng ở mỗi đầu và điều hướng bất kỳ tình huống bất ngờ nào dọc đường. Vì vậy, chúng ta không thể loại bỏ những công việc được. Mặc dù người lái xe tải là những người phù hợp và có kinh nghiệm với những công việc khác như trên, họ có thể sẽ là những người đầu tiên được tuyển dụng vào một công việc được tái định hình.

NHỮNG ĐIỂM CHÍNH

• Một công việc là sự tập hợp của nhiều nhiệm vụ. Khi phân tích luồng công việc và sử dụng các công cụ AI, một số công việc trước kia được thực hiện bởi con người có thể sẽ được tự động hóa, thứ tự và sự nhấn mạnh vào những nhiệm vụ còn lại có thể thay đổi, những nhiệm vụ mới có thể được tạo ra. Do vậy, sự tập hợp các nhiệm vụ tạo thành một công việc có thể thay đổi.

• Sự bổ sung các công cụ AI tạo ra bốn sự ảnh hưởng đến các công việc:

1. Các công cụ AI có thể giúp mở rộng công việc, như trong ví dụ về bảng tính và chuyên viên lưu trữ.

2. Các công cụ AI có thể thu gọn công việc, như trong ví dụ về hậu cần.

3. Các công cụ AI có thể dẫn đến sự tái cơ cấu công việc, với một số phần công việc được bổ sung và những phần công việc khác bị loại bỏ, như với các bác sĩ X-quang.

4. Các công cụ AI có thể làm giảm nhẹ sự nhấn mạnh vào những kỹ năng cụ thể trong một công việc cụ thể, như với những người lái xe buýt trường học.

5. Các công cụ AI có thể thay đổi lợi nhuận tương quan với những kỹ năng cụ thể và, do đó, thay đổi những kiểu người phù hợp nhất cho những công việc cụ thể. Trong trường hợp của những chuyên viên lưu trữ, sự xuất hiện của bảng tính giảm thiểu lợi ích của khả năng thực hiện nhiều phép tính nhanh chóng bằng máy tính. Đồng thời, nó tăng lợi ích của việc giỏi đặt ra câu hỏi để tận dụng tối đa khả năng phân tích viễn cảnh của công nghệ.

PHẦN 4CHIẾN LƯỢC



15Ai trong bộ máy quản lý cao cấp

V

ào tháng 1 năm 2007, khi Steve Jobs bước lên sân khấu và giới thiệu iPhone với thế giới, không ai nghĩ rằng: “Ừ thì, nó là dấu chấm hết cho nền công nghiệp taxi.” Tuy nhiên cho đến năm 2018, đó dường như lại chính xác là điều xảy ra. Trong suốt thập kỷ qua, điện thoại thông minh phát triển từ việc đơn thuần là một chiếc điện thoại thông minh hơn thành một nền tảng không thể thiếu cho các công cụ đang phá vỡ hoặc thay đổi về mặt cơ bản tất cả các hình thức của các nền công nghiệp.

Những sự phát triển gần đây của AI và máy tự học đã khiến chúng tôi tin rằng sự đổi mới này ngang hàng với những công nghệ tuyệt vời mang tính thay đổi trong quá khứ: điện, xe, nhựa, vi mạch, mạng Internet và điện thoại thông minh. Từ lịch sử kinh tế, chúng tôi biết những công nghệ đa năng này khuếch tán và thay đổi như thế nào. Chúng tôi cũng nhận ra mức độ khó khăn của việc dự đoán khi nào, nơi nào và làm thế nào mà những sự thay đổi mang tính phá vỡ này sẽ xảy ra. Đồng thời, chúng tôi đã biết cách mong chờ những gì, làm thế nào để đẩy nhanh sự thay đổi và khi nào một công nghệ mới có khả năng chuyển từ một điều gì đó thú vị sang một điều gì đó mang tính thay đổi.

Khi nào AI nên là một chương trình quan trọng trong đội ngũ lãnh đạo của tổ chức? Trong khi những sự tính toán ROI có thể ảnh hưởng đến những thay đổi trong vận hành, những quyết định chiến lược mang đến những tình huống nan giải và buộc những người lãnh đạo phải đối mặt với sự không chắc chắn. Việc sử dụng AI vào trong một phần của tổ chức có thể đòi hỏi những thay đổi ở những phần khác. Với những hiệu quả nội bộ trong tổ chức, việc sử dụng AI và những quyết định khác đòi hỏi thẩm quyền của một người có thể giám sát toàn bộ doanh nghiệp, ví dụ như CEO.

Cách mà AI có thể thay đổi chiến lược kinh doanh

Trong chương 2, chúng tôi đã phỏng đoán rằng một khi nút quay trên máy dự đoán được vặn đủ, các công ty như Amazon có thể tự tin thay đổi mô hình kinh doanh theo những nhu cầu cụ thể của khách hàng.

Đầu tiên, thế lưỡng nan chiến lược hoặc sự đánh đổi phải tồn tại. Đối với Amazon, thế lưỡng nan là mô hình vận chuyển-rồi-mua sắm có thể mang lại nhiều doanh thu hơn nhưng đồng thời mang lại nhiều sản phẩm mà khách hàng muốn đem trả lại hơn. Khi chi phí của việc trả lại sản phẩm quá cao, thì ROI cho mô hình vận chuyển- rồi-mua sắm sẽ thấp hơn ROI của cách tiếp cận truyền thống mua sắm-rồi-vận chuyển. Điều này lí giải vì sao, mô hình kinh doanh mua sắm-rồi- vận chuyển của Amazon không thay đổi.

Thứ hai, vấn đề có thể được giải quyết bằng cách giảm thiểu sự không chắc chắn. Với Amazon, đó là về nhu cầu của người tiêu dùng. Nếu bạn có thể dự đoán chính xác những gì mọi người sẽ mua, đặc biệt nếu được vận chuyển đến trước cửa nhà họ, thì bạn giảm thiểu khả năng trả hàng và gia tăng doanh thu. Việc giảm thiểu sự không chắc chắn sẽ đem lại lợi ích và giải quyết vấn đề chi phí của thế lưỡng nan.

Loại hình quản lý nhu cầu này không mới. Nó là một lý do vì sao những cửa hàng đại lý tồn tại. Những cửa hàng đại lý không thể dự đoán nhu cầu cá nhân mỗi khách hàng, nhưng họ có thể dự đoán khả năng nhu cầu từ một nhóm khách hàng. Bằng việc gộp chung những khách hàng ghé thăm một địa điểm, những cửa hàng đại lý hạn chế sự không chắc chắn về nhu cầu giữa những cá nhân khách hàng.

Thứ ba, các công ty cần một máy dự đoán có thể giảm thiểu sự không chắc chắn để thay đổi sự cân bằng trong thế lưỡng nan chiến lược. Với Amazon, một mô hình nhu cầu khách hàng chính xác có thể khiến mô hình vận chuyển-rồi-mua sắm trở nên có giá trị. Ở đây, những lợi ích của việc tăng doanh thu lớn hơn chi phí chuyển trả hàng.

Hiện giờ, nếu Amazon thực hiện mô hình này, nó sẽ tiến một bước xa hơn trong sự thay đổi mô hình doanh nghiệp. Những thay đổi này sẽ bao gồm việc đầu tư giảm thiểu chi phí trong việc đảm bảo những gói hàng bị trả lại và những dịch vụ vận chuyển để xử lý những sự trả hàng này. Cho dù thị trường vận chuyển thân thiện với khách hàng rất cạnh tranh, dịch vụ trả hàng lại là một thị trường kém phát triển hơn nhiều. Bản thân Amazon có thể thiết lập cơ sở hạ tầng xe tải đi dọc các khu vực nhà dân hằng ngày để vận chuyển hàng và nhận hàng bị trả lại, từ đó tích hợp theo chiều dọc vào mô hình trả hàng mỗi ngày. Một cách hiệu quả, Amazon có thể dịch chuyển ranh giới doanh nghiệp của họ tới ngay trước hiên nhà bạn.

Sự dịch chuyển ranh giới này đang xảy ra. Một ví dụ là công ty liên doanh thương mại điện tử của Đức có tên là Otto.[1] Một rào cản lớn đối với sự mua hàng của người tiêu dùng trên mạng Internet thay vì ở cửa hàng đại lý là những lần không chắc chắn về sự vận chuyển hàng. Nếu những người tiêu dùng có trải nghiệm giao hàng không tốt, họ sẽ không muốn quay trở lại một trang web. Otto nhận ra rằng khi việc giao hàng bị trì hoãn (lâu hơn một vài ngày), sự trả hàng sẽ tăng vọt. Những người tiêu dùng chắc chắn sẽ tìm sản phẩm tại cửa hàng đại lý trong thời gian đó và mua hàng ở đó. Ngay cả khi Otto có doanh thu, những món hàng bị trả lại vẫn cộng dồn vào chi phí của họ.

Làm thế nào để giảm thời gian giao hàng tới người tiêu dùng? Dự đoán khả năng đặt hàng của người tiêu dùng và chuẩn bị sẵn hàng ở trung tâm phân phối gần đó. Nhưng bản thân sự quản lý hàng tồn kho cũng rất tốn kém. Thay vào đó, điều bạn muốn là chỉ giữ những món hàng mà bạn có khả năng sẽ cần thôi. Để làm được việc đó, bạn sẽ muốn sự dự đoán tốt hơn về nhu cầu của khách hàng. Otto, với lượng cơ sở dữ liệu của ba tỷ giao dịch và hàng trăm những biến khác (bao gồm những kết quả tìm kiếm và thông tin nhân khẩu học), đã có thể tạo ra máy dự đoán để xử lý điều này. Giờ đây nó có thể dự đoán với độ chính xác 90% về những sản phẩm họ có thể bán ra trong vòng một tháng. Dựa vào những dự đoán này, nó đã cải thiện khâu hậu cần rất nhiều. Hàng tồn kho giảm xuống 20% và những món hàng bị trả lại hằng năm giảm xuống 2 triệu món hàng. Sự dự đoán cải thiện khâu hậu cần, từ đó giảm thiểu chi phí và gia tăng sự hài lòng của khách hàng.

Một lần nữa, chúng ta có thể thấy ba yếu tố chiến lược quan trọng. Otto gặp phải một thế lưỡng nan (làm thế nào để cải thiện thời gian vận chuyển mà không cần những kho hàng tồn kho đắt đỏ), sự không chắc chắn dẫn đến thế lưỡng nan đó (trong trường hợp này là nhu cầu chung của khách hàng ở một địa điểm) và bằng cách giải quyết sự không chắc chắn đó (ví dụ, dự đoán nhu cầu ở địa phương tốt hơn), nó có thể thiết lập nhiều cách sắp xếp hậu cần mới, yêu cầu những vị trí nhà kho mới, giao hàng trong địa phương và đảm bảo sự giao hàng tới khách hàng. Họ đã không thể đạt được tất cả những điều đó nếu không sử dụng máy dự đoán để giải quyết sự không chắc chắn quan trọng đó.

Alabama yêu dấu?

Để máy dự đoán thay đổi chiến lược của bạn, ai đó cần phải tạo ra một máy dự đoán hữu ích đối với riêng bạn. Để làm được như vậy cần phải dựa vào nhiều thứ nằm ngoài sự kiểm soát của công ty bạn.

Hãy cùng xem xét những yếu tố khiến công nghệ dự đoán trở nên có sẵn với doanh nghiệp của bạn. Để làm được điều này, chúng ta sẽ cùng đi tới những cánh đồng ngô ở Iowa vào những năm 1930. Ở đây, một vài nhà nông tiên phong đã giới thiệu một giống ngô mới mà họ tạo ra thông qua sự lai chéo trong suốt hai thập kỷ. Giống ngô lai này chuyên biệt hơn những giống ngô thông thường. Nó đòi hỏi lai chéo hai dòng ngô để cải thiện các tính năng như khả năng chịu hạn và năng suất trong môi trường địa phương cụ thể. Giống ngô lai này là một sự thay đổi quan trọng bởi vì không những nó hứa hẹn mang lại năng suất cao đáng kể, mà người nông dân còn trở nên độc lập đối với những yếu tố khác khi cần những hạt giống đặc biệt. Những hạt giống mới cần được điều chỉnh phù hợp với những điều kiện địa phương để thu được lợi nhuận cao nhất.

Như thể hiện trong hình 15-1, những người nông dân Alabama dường như lạc hậu hơn so với những người nông dân Iowa. Nhưng khi chuyên gia kinh tế ở Harvard Zvi Griliches nhìn kĩ vào những con số, ông phát hiện rằng khoảng cách 20 năm giữa sự áp dụng ở Iowa và Alabama không phải vì những người nông dân ở Alabama chậm, mà vì chỉ số ROI cho giống ngô lai ở những trang trại Alabama không thể hiện rõ sản lượng vào những năm 1930.[2] Những trang trại Alabama nhỏ hơn, với tỷ suất lợi nhuận thấp hơn so với những nơi ở phía Bắc và phía Tây. Ngược lại, những người nông dân Iowa có thể áp dụng giống ngô này khắp những trang trại rộng lớn và thu được lợi nhuận cao hơn để thể hiện những chi phí giống cao hơn. Một trang trại lớn đồng nghĩa với việc thử nghiệm các giống lai mới dễ dàng hơn vì người nông dân chỉ cần dành một khoảng đất nhỏ để thực hiện cho tới khi các giống lai chứng minh là có hiệu quả.[3] Những rủi ro của người nông dân Iowa thấp hơn và họ có tỷ số lợi nhuận cao hơn để đề phòng.



Trong thế giới AI, Iowa chính là Google. Công ty có hơn hàng nghìn dự án phát triển công cụ AI đang được thực hiện khắp các danh mục kinh doanh, từ việc tìm kiếm cho đến những biển quảng cáo, bản đồ và dịch thuật.[4] Những

gã khổng lồ công nghệ khác trên thế giới đã tham gia cùng Google. Lý do cũng khá rõ ràng: Google, Facebook, Baidu, Alibaba, Salesforce và những công ty khác đã có những công cụ kinh doanh trong tay. Họ có những công việc được xác định rõ ràng khiến họ mở rộng doanh nghiệp và trong mỗi công việc, đôi khi AI có thể cải thiện yếu tố dự đoán một cách đáng kể.

Những công ty lớn có tỷ số lợi nhuận cao, vậy nên họ có đủ khả năng để thử nghiệm. Họ có thể lấy một phần "đất" và sử dụng nó để thử nhiều loại hình AI mới. Họ có thể thu được lợi nhuận lớn từ những thử nghiệm thành công như vậy thông qua việc áp dụng vào hàng loạt sản phẩm hoạt động trên quy mô lớn.

Với nhiều những doanh nghiệp khác, con đường dẫn đến AI ít rõ ràng hơn. Không giống như Google, nhiều công ty không đầu tư đủ hai thập kỷ vào việc kỹ thuật số hóa tất cả các khía cạnh của luồng công việc và cũng không có một khái niệm rõ ràng về điều mà họ muốn dự đoán. Nhưng một khi công ty đã đặt ra những chiến lược được xác định rõ ràng, họ có thể phát triển những thành phần đó và đặt nền tảng cho AI hiệu quả.

Khi những điều kiện đó đúng, tất cả những người nông dân trồng ngô ở Wisconsin, Kentucky, Texas và Alabama đều làm theo những người nông dân Iowa trong việc áp dụng giống ngô lai. Lợi ích của nhu cầu đủ lớn, và chi phí của việc cung cấp giảm. Tương tự, giá thành và những rủi ro liên quan đến AI sẽ giảm đi theo thời gian, vậy nên nhiều doanh nghiệp tuy không đi đầu việc phát triển các công cụ kỹ thuật số vẫn sẽ áp dụng nó. Khi làm như vậy, nhu cầu sẽ thúc đẩy họ: cơ hội để giải quyết những thế lưỡng nan cơ bản trong mô hình doanh nghiệp bằng cách giảm thiểu sự không chắc chắn.

Bổ sung những cầu thủ bóng chày

Chiến lược của Billy Beane trong *Moneyball* – sử dụng dự đoán thống kê số liệu để vượt qua những thành kiến của những huấn luyện viên bóng chày và cải thiện sự dự đoán – là một ví dụ của việc sử dụng sự dự đoán để giảm thiểu sự không chắc chắn và cải thiện hiệu suất của Oakland Athlectics. Đó cũng là một sự thay đổi chiến lược yêu cầu sự sửa đổi hệ thống cấp bậc rõ ràng và không rõ ràng của tổ chức.

Sự dự đoán tốt hơn thay đổi ai là người mà đội nên tuyển trên sân, nhưng sự

hoạt động của đội bóng chày thì giữ nguyên không thay đổi. Các cầu thủ mà máy dự đoán lựa chọn đóng vai trò giống như các cầu thủ nó đã thay thế, có lẽ chỉ với một vài lượt đi. Và các huấn luyện viên tiếp tục đóng vai trò trong việc lựa chọn cầu thủ.[5]

Càng nhiều thay đổi cơ bản xuất hiện trong đội được tuyển chọn để chơi trên sân bóng thì kết quả sẽ là sự tái cấu trúc của biểu đồ tổ chức. Quan trọng nhất, đội tuyển chọn những cầu thủ có thể nói cho máy cần dự đoán điều gì và rồi sử dụng những sự dự đoán đó để xác định xem nên thuê những cầu thủ nào. Đội cũng đã tạo ra một chức năng công việc mới, được gọi là "chuyên gia toán học bóng chày". Một chuyên gia toán học bóng chày phát triển những biện pháp khác nhau cho những thành tựu mà đội sẽ đạt được từ việc ký hợp đồng với những cầu thủ khác nhau. Các chuyên gia toán học bóng chày là REF của bóng chày. Hiện giờ, đa số các đội đều có ít nhất một chuyên gia như vậy và vai trò như vậy đã xuất hiện, dưới những cái tên khác nhau, trong những môn thể thao khác.

Sự dự đoán tốt hơn tạo ra một vị trí cao cấp hơn trên biểu đồ của tổ chức. Các nhà khoa học nghiên cứu, nhà khoa học dữ liệu và phó chủ tịch phân tích dữ liệu được liệt kê như những người đóng vai trò quan trọng trong các thư mục văn phòng trực tuyến. Houston Astros thậm chí có cả một đơn vị khoa học đưa ra quyết định độc lập đứng đầu bởi người từng là kỹ sư của NASA, Sig Mejdal. Sự thay đổi chiến lược cũng đồng nghĩa với sự thay đổi người mà đội tuyển chọn để lựa chọn các cầu thủ. Những chuyên gia phân tích sở hữu các kỹ năng toán học, nhưng chỉ những người giỏi nhất trong số họ hiểu rõ cần phải nói gì để máy dự đoán hoạt động. Họ cung cấp sự đánh giá.

Sự lựa chọn chiến lược yêu cầu sự đánh giá mới

Sự thay đổi trong tổ chức của quản lý đội bóng chày nhấn mạnh một vấn đề khác cho bộ máy quản lý cao cấp trong việc bổ sung những lựa chọn chiến lược liên quan đến AI. Trước khi có toán học bóng chày, sự đánh giá của các huấn luyện viên bóng chày bị hạn chế bởi những nhược và ưu điểm của cá nhân các cầu thủ. Nhưng việc sử dụng những sự đo lường định lượng có thể dự đoán cách chơi hiệu quả của nhóm cầu thủ. Sự đánh giá chuyển từ suy nghĩ về sự trả giá của một cầu thủ cụ thể sang suy nghĩ về sự trả giá của một đội cụ thể. Sự dự đoán tốt hơn giờ đây cho phép người quản lý đưa ra những

quyết định gần với mục tiêu của tổ chức hơn: xác định đội tốt nhất thay vì xác định những cá nhân cầu thủ tốt nhất.

Để tận dụng tối đa máy dự đoán, bạn cần phải suy nghĩ lại về những chức năng phần thưởng trong tổ chức của bạn để phù hợp hơn với những mục tiêu thực sự. Việc này không hề dễ chút nào. Ngoài việc tuyển dụng, đội cũng cần thay đổi chiến lược marketing, có lẽ để giảm bớt sự tập trung vào hiệu suất cá nhân. Tương tự, các huấn luyện viên cần phải hiểu lý do tuyển những cầu thủ đơn lẻ và tầm ảnh hưởng của việc sắp xếp vị trí các cầu thủ trong mỗi trận đấu. Cuối cùng, ngay cả những cầu thủ cũng cần hiểu vai trò của họ có thể thay đổi phụ thuộc vào việc liệu những đối thủ của họ có sử dụng các công cụ dự đoán tương tự hay không.

Những lợi thế có thể bạn đã có

Chiến lược cũng có nghĩa nắm bắt giá trị - tức là, ai sẽ là người nắm bắt giá trị mà sự dự đoán tốt hơn tạo ra?

Các giám đốc điều hành kinh doanh thường nói với chúng ta rằng bởi vì máy dự đoán cần dữ liệu, bản thân dữ liệu cũng là một loại tài sản chiến lược. Đó là, nếu bạn có nhiều năm thu thập dữ liệu, ví dụ, doanh số bán sữa chua, thì để dự đoán được doanh số bán sữa chua bằng máy dự đoán, ai đó sẽ cần loại dữ liệu đó. Do đó, nó có giá trị với người sở hữu. Nó giống như việc có một kho trữ dầu vậy.

Giả định đó che đi một vấn đề quan trọng – giống như dầu, dữ liệu cũng có nhiều lớp khác nhau. Chúng tôi đã nhấn mạnh ba loại dữ liệu – đào tạo, đầu vào, và phản hồi. Dữ liệu đào tạo được sử dụng để xây dựng máy dự đoán. Dữ liệu đầu vào được sử dụng để tạo ra sự dự đoán. Dữ liệu phản hồi được sử dụng để cải thiện nó. Chỉ hai loại dữ liệu được đề cập sau là cần cho việc sử dụng trong tương lai. Dữ liệu đào tạo được sử dụng lúc đầu để đào tạo một thuật toán, nhưng một khi máy dự đoán bắt đầu chạy, nó không còn hữu ích nữa. Giống như là bạn đã đốt cháy nó. Dữ liệu trong quá khứ của bạn về doanh số sữa chua có ít giá trị hơn một khi bạn sở hữu một máy dự đoán được xây dựng từ thông tin đó.[6] Nói cách khác, bây giờ nó có thể có giá trị, nhưng nó không có khả năng trở thành một nguồn giá trị bền vững. Để làm được điều đó bạn cần phải hoặc tạo ra dữ liệu mới – cho đầu vào hoặc phản hồi – hoặc bạn cần một lợi thế khác. Chúng tôi sẽ khám phá những lợi thế của

việc tạo ra dữ liệu mới trong chương tiếp theo và tập trung vào những lợi thế khác bây giờ.

Dan Bricklin, người phát minh ra bảng tính, đã tạo ra giá trị to lớn, nhưng ông không phải là người giàu có. Vậy giá trị của bảng tính đã đi đâu hết? Trên bảng xếp hạng giàu có, những người bắt chước như người sáng lập Lotus 1-2-3, Mitch Kapor hay Bill Gates của Microsoft chắc chắn vượt xa Bricklin, nhưng ngay cả họ cũng chỉ là một phần nhỏ của giá trị bảng tính. Thay vào đó, giá trị được chuyển đến những người dùng, những doanh nghiệp sử dụng bảng tính để làm ra những quyết định tốt hơn trị giá tỷ đô. Cho dù Lotus và Microsoft có làm gì, những người dùng của họ sẽ quyết định bảng tính đã được cải thiện hay chưa.

Bởi vì họ hoạt động ở cấp độ quyết định, điều này cũng đúng với máy dự đoán. Hãy tưởng tượng những ứng dụng AI hỗ trợ trong việc quản lý hàng tồn kho cho một chuỗi siêu thị. Việc biết khi nào lượng sữa chua sẽ được bán giúp bạn biết khi nào bạn nên dự trữ nó và giảm thiểu lượng sữa chua chưa bán phải bỏ đi. Một người sáng tạo AI mà cung cấp máy dự đoán về nhu cầu sữa chua có thể sẽ thu lợi, nhưng sẽ phải kết hợp với một chuỗi siêu thị để tạo ra bất kỳ giá trị nào. Chỉ chuỗi siêu thị mới có thể thực hiện hành động dự trữ sữa chua hay không. Và nếu không có hành động đó, máy dự đoán cho nhu cầu sữa chua sẽ không có giá trị. Nhiều doanh nghiệp sẽ tiếp tục hành động khi có hoặc không có AI. Họ sẽ có lợi thế hơn trong việc nắm bắt một số giá trị phát sinh từ việc áp dụng AI. Lợi thế này không đồng nghĩa với việc các công ty tự hành động sẽ nắm bắt được tất cả mọi giá trị.

Trước khi bán bảng tính, Bricklin và đối tác của ông, Bob Frankston, tự hỏi rằng liệu họ có nên giữ nó không. Sau đó họ có thể bán những mô hình kỹ năng của họ và, kết quả là, nắm bắt được giá trị được tạo ra bởi những hiểu biết sâu sắc của họ. Họ từ bỏ kế hoạch này - có thể là vì lý do chính đáng—nhưng trong lĩnh vực AI, chiến lược của họ có thể hiệu quả. Các nhà cung cấp dịch vụ AI có thể cố gắng cản trở các công ty truyền thống.

Ở một mức độ nào đó, các phương tiện tự động chính là một ví dụ. Trong khi một số các nhà sản xuất ô tô truyền thống đang đầu tư mạnh mẽ vào khả năng của chính họ, những người khác đang hy vọng sẽ hợp tác với những người ngoài ngành (ví dụ như Waymo của Alphabet) thay vì phát triển các khả năng đó trong ngành. Trong các trường hợp khác, các công ty công nghệ lớn đang

bắt đầu triển khai các dự án với các nhà sản xuất ô tô truyền thống. Ví dụ, Baidu, nhà điều hành công cụ tìm kiếm lớn nhất của Trung Quốc, đang dẫn đầu một sáng kiến lái xe tự động mở rộng và đa dạng với Apollo Project, cùng nhiều đối tác khác, bao gồm Daimler và Ford. Chen Juhong, phó chủ tịch của Tencent, nhận xét rằng, “Tencent hy vọng có thể nỗ lực hết sức để củng cố sự phát triển của các công nghệ AI được sử dụng trong lái xe tự động… Chúng tôi muốn trở thành ‘người kết nối’ để hỗ trợ thúc đẩy sự hợp tác, sự đổi mới và sự kết hợp trong ngành…”[7]. Suy nghĩ về áp lực cạnh tranh thúc đẩy sự hợp tác, Chủ tịch Xu Heyi của Beijing Automotive nói rằng: “Trong kỷ nguyên mới này, chỉ những công ty kết nối với các công ty khác để xây dựng thế hệ xe tiếp theo mới có thể tồn tại, còn những người đóng cửa để một mình tạo ra các phương tiện thì sẽ không tồn tại được.”[8]

Những đối tượng tương đối mới (ví dụ như Tesla) đang cạnh tranh với những cái tên nổi bật bằng cách triển khai trực tiếp AI trên những chiếc xe mới, tích hợp chặt chẽ phần mềm và phần cứng. Trong ngành công nghiệp đó, cuộc đua cho việc nắm bắt giá trị không tôn trọng những ranh giới kinh doanh truyền thống.

Nguyên tắc kinh tế đơn giản của chiến lược AI

Những thay đổi chúng tôi đã nhấn mạnh phụ thuộc vào hai khía cạnh khác nhau của AI tác động lên cốt lõi khuôn mẫu kinh tế của chúng ta.

Đầu tiên, như trong mô hình vận chuyển-rồi-mua sắm của Amazon, máy dự đoán giảm thiểu sự không chắc chắn. Khi AI tiến bộ, chúng ta sẽ sử dụng máy dự đoán để giảm thiểu sự không chắc chắn một cách bao quát hơn. Do đó, những thế lưỡng nan chiến lược được thúc đẩy bởi sự không chắc chắn sẽ tiến hóa cùng với AI. Khi giá thành của AI giảm, máy dự đoán sẽ giải quyết được nhiều thế lưỡng nan chiến lược hơn.

Thứ hai, AI sẽ tăng giá trị của những yếu tố bổ sung cho sự dự đoán. Đánh giá của một nhà phân tích bóng chày, hành động của một nhà tạp hóa bán lẻ và như chúng tôi sẽ chỉ ra trong chương 17 – dữ liệu của máy dự đoán trở nên quan trọng đến mức bạn có thể sẽ phải thay đổi chiến lược của mình để tận dụng những gì nó cung cấp.

NHỮNG ĐIỂM CHÍNH

• Bộ máy quản lý cao cấp không được ủy quyền toàn bộ chiến lược AI cho bộ phận IT vì các công cụ AI mạnh mẽ có thể nâng cao năng suất của các công việc hơn rất nhiều so với chiến lược của tổ chức và từ đó dẫn đến sự thay đổi bản chất chiến lược. AI có thể dẫn đến sự thay đổi chiến lược nếu có ba yếu tố: (1) có một sự đánh đổi cốt lõi trong mô hình kinh doanh (ví dụ, mua sắm-rồi-vận chuyển với vận chuyển-rồi- mua sắm); (2) sự đánh đổi bị ảnh hưởng bởi sự không chắc chắn (ví dụ, doanh thu cao hơn từ mô hình vận chuyển-rồi-mua sắm bị đánh bại bởi chi phí của các mặt hàng bị trả lại do sự không chắc chắn về những gì khách hàng sẽ mua); và (3) một công cụ AI giảm thiểu sự không chắc chắn sẽ thay đổi cán cân của sự đánh đổi và khiến chiến lược tối ưu thay đổi từ phía này sang phía khác (ví dụ, AI làm giảm sự không chắc chắn bằng cách dự đoán nhu cầu của khách hàng sẽ thay đổi cán cân, khiến sự trả lại hàng từ mô hình vận chuyển-rồi-mua sắm trở nên lợi hơn mô hình truyền thống).

• Một lý khác khiến bộ phận quản lý cao cấp cần cho chiến lược AI là bởi sự bổ sung các công cụ AI trong một phần của doanh nghiệp cũng có thể ảnh hưởng đến các bộ phận khác. Trong thử nghiệm của Amazon, một tác dụng phụ của việc chuyển sang mô hình vận chuyển-rồi-mua sắm là sự tích hợp theo chiều dọc vào việc thu thập các mặt hàng bị trả lại, có lẽ với hàng loạt xe tải thực hiện nhận đồ hằng tuần khắp các khu phố. Nói cách khác, các công cụ AI mạnh mẽ có thể dẫn đến sự tái thiết kế đáng kể các luồng công việc và ranh giới của công ty.

• Máy dự đoán sẽ tăng giá trị của những yếu tố bổ sung, bao gồm sự đánh giá, hành động và dữ liệu. Việc giá trị của sự đánh giá ngày càng tăng có thể dẫn đến những sự thay đổi trong hệ thống phân cấp tổ chức – tức là bạn sẽ thu được nhiều lợi ích hơn khi đặt những vai trò khác nhau hoặc những người khác nhau ở các vị trí quyền lực. Ngoài ra, máy dự đoán cho phép những người quản lý vượt xa việc tối ưu hóa những thành phần đơn lẻ tới việc tối ưu hóa những mục tiêu cao cấp hơn và từ đó khiến các quyết định trở nên gần hơn với mục tiêu của tổ chức. Việc hành động dưới sự ảnh hưởng của sự dự đoán có thể là một nguồn cạnh tranh lợi thế cho phép các doanh nghiệp truyền thống nắm bắt một số giá trị từ AI. Tuy nhiên, trong một số trường hợp, các công cụ AI mạnh mẽ cung cấp một lợi thế cạnh tranh đáng kể, những đối tượng mới tham gia có thể tích hợp theo chiều dọc vào việc sở hữu

hành động và tận dụng AI của họ làm cơ sở cho sự cạnh tranh.

16Khi AI thay đổi doanh nghiệp của bạn

J

oshua (một trong những tác giả) gần đây đã hỏi một công ty máy tự học ở giai đoạn đầu, “Tại sao các bạn lại chọn cung cấp cho bác sĩ sự chẩn đoán?” Công ty của họ đầu tư vào việc xây dựng một công cụ AI có khả năng nói cho bác sĩ biết liệu một tình huống y tế cụ thể đã xuất hiện hay chưa. Một nhị phân dữ liệu đầu ra đơn giản. Một chẩn đoán. Vấn đề là, để có thể làm được điều đó, công ty phải có được sự chấp thuận theo quy định, điều này đòi hỏi những thử nghiệm tốn kém. Để quản lý những thử nghiệm đó, họ đang xem xét liệu có nên hợp tác với một công ty dược phẩm có uy tín hoặc một công ty thiết bị y tế hay không. Câu hỏi của Joshua liên quan đến chiến lược hơn là y tế. Tại sao họ lại muốn cung cấp sự chẩn đoán cho bác sĩ? Thay vào đó, tại sao họ không đơn thuần cung cấp sự dự đoán?

Joshua gợi ý rằng công ty nên tập trung vào sự dự đoán thay vì chẩn đoán. Ranh giới của doanh nghiệp họ có thể kết thúc với sự dự đoán. Điều này làm giảm bớt nhu cầu phê duyệt theo quy định, bởi vì các bác sĩ có nhiều công cụ để đi đến kết luận chẩn đoán. Công ty không cần phải hợp tác ở giai đoạn đầu với các công ty có uy tín. Quan trọng nhất, họ không cần phải nghiên cứu và tìm ra cách để diễn giải sự dự đoán thành sự chẩn đoán. Tất cả những gì họ phải suy luận chỉ là độ chính xác cần thiết để cung cấp một sự dự đoán có giá trị. Là 70, 80, hay 99%?

Doanh nghiệp của bạn sẽ kết thúc ở đâu và các doanh nghiệp khác bắt đầu từ đâu? Ranh giới chính xác của công ty bạn ở đâu? Quyết định về mặt lâu dài này đòi hỏi sự chú ý cẩn thận từ cấp cao nhất của tổ chức. Hơn nữa, những đổi mới đa năng thường dẫn đến các câu trả lời mới cho câu hỏi về ranh giới. Một số công cụ AI cụ thể có khả năng thay đổi những ranh giới của doanh nghiệp bạn. Máy dự đoán sẽ thay đổi cách các doanh nghiệp suy nghĩ về tất cả mọi thứ, từ thiết bị vốn của họ đến dữ liệu và con người của họ.

Điều gì nên giữ lại và điều gì nên loại bỏ

Sự không chắc chắn có thể gây ảnh hưởng đến những ranh giới của một doanh nghiệp.[1] Các chuyên gia kinh tế Silke Forbes và Mara Lederman đã đánh giá tổ chức của Ngành công nghiệp hàng không Hoa Kỳ ở bước chuyển sang thời kì thiên niên kỷ.[2] Những hãng hàng không lớn như của United và American đã khai thác một số tuyến đường, trong khi đó những đối tác khu vực như American Eagle và SkyWest xử lý những tuyến khác. Các đối tác là các doanh nghiệp độc lập có thỏa thuận hợp đồng với hãng lớn. Không có những sự cân nhắc khác, các hãng hàng không khu vực thường hoạt động với chi phí thấp hơn so với các hãng lớn, tiết kiệm tiền lương và ít các quy tắc làm việc có lợi hơn. Ví dụ, một số nghiên cứu cho thấy các phi công cao cấp tại các hãng lớn nhận được mức lương cao hơn 80% so với những phi công ở các hãng hàng không khu vực.

Câu hỏi là lý do tại sao các hãng lớn mà không phải là những đối tác khu vực khai thác nhiều tuyến đường như vậy, cho dù các đối tác có thể cung cấp dịch vụ ở mức chi phí thấp hơn. Forbes và Lederman xác định một yếu tố thúc đẩy - thời tiết - hoặc cụ thể hơn, sự không chắc chắn về thời tiết. Khi một sự kiện thời tiết xảy ra bất thường, nó sẽ trì hoãn các chuyến bay, trong ngành công nghiệp hàng không được quản lý chặt chẽ và liên kết chặt chẽ, điều này có thể tạo nên hiệu ứng gợn sóng trong toàn bộ hệ thống. Khi thời tiết trở xấu, các hãng hàng không lớn không muốn bị ngăn cản bởi các đối tác khi họ phải thực hiện những sự thay đổi nhanh chóng với chi phí không biết chắc. Vì vậy, đối với các tuyến đường có thể có sự chậm trễ liên quan đến thời tiết, đa số các hãng hàng không lớn giữ quyền kiểm soát và hoạt động.

Ba thành phần chúng tôi đã nhấn mạnh trong chương trước gợi ý rằng AI có thể dẫn đến sự thay đổi chiến lược. Đầu tiên, chi phí thấp hơn so với nhiều sự kiểm soát hơn là một sự đánh đổi cốt lõi. Thứ hai, sự đánh đổi đó được gây ra gián tiếp bởi sự không chắc chắn; đặc biệt, sự gia tăng kiểm soát cũng đi cùng với mức độ không chắc chắn. Các hãng hàng không lớn cân bằng chi phí thấp hơn và nhiều sự kiểm soát hơn bằng cách tối ưu hóa những ranh giới nơi những hoạt động của họ kết thúc và những hoạt động mà những đối tác của họ bắt đầu. Nếu máy dự đoán có thể cắt qua sự không chắc chắn này, thì thành phần thứ ba sẽ có mặt và sự cân bằng sẽ thay đổi. Các hãng hàng không sẽ ký nhiều hợp đồng hơn với các đối tác của họ.

Các doanh nghiệp tham gia vào sự đổi mới liên tục, đặc biệt là sự đổi mới

liên quan đến việc học hỏi từ kinh nghiệm, tạo ra một mẫu tương tự. Trong suốt quá trình phát triển lâu dài, một nhà sản xuất ô tô chỉ có những hiểu biết hữu hạn về cách hoạt động của một mô hình mới. Một số thông tin chỉ có thể được thu thập sau khi sản phẩm ra mắt, ví dụ như phản hồi của khách hàng và những phép đo hiệu suất dài hạn khác. Đây là lý do chính tại sao các mô hình được cập nhật hằng năm với thiết kế không thay đổi nhưng cung cấp những cải tiến cho các thành phần hoạt động và cải thiện sản phẩm.

Các chuyên gia kinh tế Sharon Novak và Scott Stern nhận ra rằng các nhà sản xuất xe ô tô tự động hạng sang mà tự sản xuất các bộ phận riêng phát triển nhanh hơn cho dù có sự thay đổi mô hình qua các năm.[3] Họ đo lường được những sự cải thiện từ phía khách hàng bằng việc sử dụng xếp hạng từ Báo cáo người tiêu dùng (consumer report). Việc có quyền kiểm soát đồng nghĩa với việc các nhà sản xuất ô tô tự động có thể sẵn sàng thích nghi hơn với phản hồi của khách hàng.

Ngược lại, những công ty thu mua các bộ phận không cho thấy sự cải thiện tương tự. Tuy nhiên, những công ty này nhận được một lợi ích khác; những mô hình ban đầu của họ có chất lượng cao hơn so với những mô hình đầu tiên của các nhà sản xuất ô tô tự chế tạo các bộ phận riêng. Mô hình mới mà ở đó các nhà sản xuất thu mua các bộ phận tốt hơn ngay từ đầu bởi các nhà cung cấp chế tạo bộ phận tốt hơn. Do vậy, các nhà sản xuất ô tô đối mặt với lựa chọn thu mua hoặc tự chế tạo các bộ phận để đạt được cải thiện theo thời gian khi họ kiểm soát sự đổi mới trong chu kỳ của mô hình sản phẩm.

Trong mỗi trường hợp, sự đánh đổi giữa hiệu suất ngắn hạn và dài hạn cũng như những sự kiện theo thói quen và không theo thói quen đều được giải quyết bằng một lựa chọn tổ chức quan trọng: nên phụ thuộc vào các nhà cung cấp bên ngoài đến mức nào. Nhưng sự nổi bật của lựa chọn đó liên quan chặt chẽ đến sự không chắc chắn. Các sự kiện thời tiết quan trọng đến mức nào mà khiến các hãng hàng không không thể lên kế hoạch trước? Làm sao để biết chiếc xe có phù hợp với những gì khách hàng thực sự muốn?

Tác động của AI: Nguồn vốn

Hãy giả sử có một AI có thể làm giảm thiểu sự không chắc chắn này, vậy là có sự xuất hiện của thành phần thứ ba. Sự dự đoán có giá thành rẻ đến mức nó tối giản sự không chắc chắn đủ để thay đổi bản chất của thế lưỡng nan

chiến lược. Điều này sẽ ảnh hưởng như thế nào đến các hãng hàng không và các nhà sản xuất ô tô? AI có thể cho phép máy hoạt động trong những môi trường phức tạp hơn. Nó gia tăng số lượng “nếu” đáng tin cậy, từ đó làm giảm nhu cầu của doanh nghiệp sở hữu nguồn thiết bị của riêng mình, vì hai lý do.

Đầu tiên, nhiều “nếu” đồng nghĩa một doanh nghiệp có thể soạn thảo các hợp đồng để xác định phải làm gì khi có điều gì đó bất thường xảy ra. Giả sử AI cho phép các hãng hàng không không những có thể dự báo các sự kiện thời tiết mà còn có thể tạo ra các dự đoán tốt nhất để giải quyết các gián đoạn liên quan đến thời tiết. Điều này sẽ tăng lợi ích cho các hãng hàng không lớn vì đã cụ thể hơn trong hợp đồng để xử lý với những phát sinh. Họ có thể xác định số lượng “nếu” nhiều hơn trong hợp đồng. Do đó, thay vì kiểm soát các tuyến đường hàng không thông qua quyền sở hữu, các hãng hàng không lớn sẽ có khả năng dự đoán để tự tin hơn khi soạn thảo hợp đồng với các hãng hàng không hoạt động độc lập trong khu vực, cho phép họ tận dụng lợi thế của chi phí thấp hơn. Họ sẽ yêu cầu ít nguồn thiết bị hơn (ví dụ như máy bay), bởi vì họ có thể thu mua nhiều chuyến bay từ các hãng vận tải bé hơn trong khu vực.

Thứ hai, sự dự đoán dựa trên AI - dự đoán mức độ hài lòng của người tiêu dùng - sẽ cho phép các nhà sản xuất ô tô tự tin hơn khi thiết kế những sản phẩm, từ đó dẫn đến mức độ hài lòng của người tiêu dùng cao và hiệu suất tốt mà không có nhu cầu chỉnh sửa bao quát ở giữa mô hình. Nhờ đó, các nhà sản xuất ô tô sẽ có thể chọn phần tốt nhất trên thế giới cho các mô hình của họ từ các nhà cung cấp độc lập, tự tin rằng sự dự đoán vượt trội đã loại bỏ sự cần thiết của việc tái đàm phán hợp đồng tốn kém.

Tác động của AI: Lao động

Các ngân hàng tung ra máy rút tiền tự động (ATM) được phát triển trong những năm 1970 và sử dụng rộng rãi trong suốt những năm 1980. Khả năng tiết kiệm lao động của công nghệ này là - như tên gọi của nó - được thiết kế để tự động hóa các nhân viên giao dịch.

Theo Cục Thống kê Lao động, các nhân viên giao dịch không bị mất việc (xem hình 16-1). Họ đã trở thành người tiếp thị và là người đại diện dịch vụ khách hàng cho các sản phẩm ngân hàng ngoài việc thu thập và phân chia

tiền mặt. Một lý do khiến các ngân hàng không muốn mở thêm nhiều chi nhánh là do vấn đề an ninh và chi phí của con người dành quá nhiều thời gian vào một việc gì đó ví dụ như giao dịch ngân hàng. Được giải phóng khỏi những hạn chế đó, các chi nhánh ngân hàng tăng trưởng nhanh (hơn 43% ở các khu đô thị), về hình dạng và kích cỡ, và với họ, một nhân viên sẽ được gọi là “nhân viên giao dịch”.



Sự ra đời của các máy ATM tạo ra một sự thay đổi đáng kể trong tổ chức; nhân viên giao dịch mới được yêu cầu phải có nhiều sự đánh giá khách quan hơn. Theo định nghĩa, các công việc trước đây của nhân viên giao dịch là theo trình tự và dễ dàng bị cơ giới hóa. Nhưng những công việc mới bao gồm việc nói chuyện với khách hàng về những nhu cầu ngân hàng của họ, tư vấn cho họ về các khoản vay và xác định những lựa chọn thẻ tín dụng đều trở nên phức tạp hơn.

Trong quá trình, việc đánh giá liệu các nhân viên giao dịch mới đã làm tốt hay không trở nên khó khăn hơn.[4]

Khi các biện pháp đo lường hiệu suất thay đổi từ khách quan (có phải bạn đang giữ ngân phiếu ngân hàng?) thành chủ quan (bạn có đang bán đúng loại sản phẩm?), quản lý nguồn nhân lực (HR) trở nên phức tạp hơn. Các chuyên gia kinh tế sẽ cho bạn biết rằng những trách nhiệm công việc có thể sẽ phải trở nên ít rõ ràng và có nhiều mối liên quan hơn. Bạn sẽ đánh giá và khen thưởng nhân viên dựa trên các quy trình chủ quan, ví dụ như đánh giá hiệu suất trong khi tính đến sự phức tạp của các công việc và các điểm mạnh cũng như điểm yếu của các nhân viên. Các quá trình như vậy rất khó để thực hiện vì việc phụ thuộc vào chúng để tạo ra những động lực cho hiệu suất tốt đòi hỏi rất nhiều niềm tin. Tuy nhiên, khi các biện pháp đo lường hiệu suất là chủ quan trong những môi trường phức tạp, những lỗi nghiêm trọng có thể xảy ra, ví dụ như ngân hàng Wells Fargo với sự gian lận của những người quản lý khách hàng.[5]

Ảnh hưởng trực tiếp của dòng logic kinh tế này là AI sẽ thay đổi quản lý nhân sự theo hướng liên quan đến các mối quan hệ và tránh xa những sự giao dịch. Có ít nhất là hai lý do. Đầu tiên, sự đánh giá của con người sẽ được tận

dụng ở những nơi có giá trị bởi vì rất khó để lập trình sự đánh giá như vậy vào máy. Thứ hai, trong trường hợp đánh giá của con người trở nên quan trọng hơn khi những dự đoán của máy xuất hiện nhiều hơn, nó sẽ bao gồm những phương tiện đánh giá hiệu suất chủ quan. Nếu những phương tiện khách quan có sẵn, một máy rất có thể đưa ra sự đánh giá như vậy mà không cần bất kỳ đánh giá nào của con người. Vì vậy, con người rất quan trọng đối với quá trình đưa ra quyết định để khiến các mục tiêu trở thành chủ quan.

Các lực lượng ảnh hưởng đến nguồn thiết bị cũng ảnh hưởng đến lao động. Nếu thông tin đầu ra quan trọng của lao động là dữ liệu, sự dự đoán hoặc hành động, vậy việc sử dụng AI có nghĩa là có nhiều hợp đồng lao động thu mua hơn, đồng thời nhiều thiết bị và vật tư thu mua hơn. Với vốn, sự dự đoán tốt hơn cung cấp thêm nhiều "nếu" mà chúng ta có thể sử dụng để xác định rõ ràng "thì" trong một hợp đồng thu mua.

Tuy nhiên, sự ảnh hưởng quan trọng hơn đến lao động sẽ là sự gia tăng tầm quan trọng trong đánh giá của con người. Sự dự đoán và sự đánh giá là những yếu tố bổ sung, vậy nên sự dự đoán tốt hơn gia tăng nhu cầu của sự đánh giá, điều này có nghĩa vai trò chính của các nhân viên sẽ là thực hiện sự đánh giá trong quá trình đưa ra quyết định. Điều này, theo định nghĩa, không thể được xác định chính xác trong một hợp đồng. Ở đây, máy dự đoán làm gia tăng sự không chắc chắn trong thế lưỡng nan chiến lược bởi vì đánh giá chất lượng của sự đánh giá là rất khó, vậy nên việc ký kết hợp đồng sẽ nguy hiểm. Ngược lại, sự dự đoán tốt hơn gia tăng sự không chắc chắn về chất lượng công việc của con người: bạn cần phải giữ những nhân viên khác trong nội bộ tập trung vào việc phán đoán.

Tác động của AI: Dữ liệu

Một vấn đề chiến lược quan trọng khác là quyền sở hữu và kiểm soát dữ liệu. Trong khi những tác động của AI đến các công nhân liên quan đến sự bổ trợ giữa sự dự đoán và sự đánh giá, mối quan hệ giữa sự dự đoán và dữ liệu cũng thúc đẩy những sự đánh đổi này. Dữ liệu giúp sự dự đoán tốt hơn. Ở đây, chúng tôi xem xét những sự đánh đổi liên quan đến những ranh giới của tổ chức. Bạn nên sử dụng dữ liệu của người khác hay sở hữu dữ liệu của riêng mình? (Trong chương tiếp theo, chúng tôi khám phá các vấn đề liên quan đến tầm quan trọng chiến lược của việc đầu tư vào thu thập dữ liệu).

Đối với các công ty khởi nghiệp AI, việc sở hữu dữ liệu cho phép họ học hỏi là một điều rất quan trọng. Nếu không, họ sẽ không thể cải thiện sản phẩm của họ theo thời gian. Công ty khởi nghiệp về máy tự học Ada Support giúp đỡ các công ty khác tương tác với khách hàng của họ. Ada có cơ hội để tích hợp sản phẩm của mình vào hệ thống của một nhà cung cấp có uy tín. Nếu việc này thành công, nó sẽ dễ dàng thu hút nhiều người hơn và thiết lập một cơ sở người dùng lớn hơn. Đây là một cách hấp dẫn để đi theo. Tuy nhiên, vấn đề là các công ty có uy tín sẽ sở hữu dữ liệu phản hồi về sự tương tác. Không có dữ liệu đó, Ada sẽ không thể cải thiện sản phẩm của mình dựa trên những gì thực sự xảy ra trong ngành. Ada được khuyến khích xem xét lại cách tiếp cận này và không tích hợp cho đến khi họ có thể đảm bảo rằng họ sở hữu dữ liệu kết quả. Làm như vậy đã cho công ty một kênh cung cấp dữ liệu bây giờ và trong tương lai để có thể học hỏi liên tục.

Vấn đề về việc sở hữu hay mua dữ liệu là quá sức với các công ty khởi nghiệp. Dữ liệu được thiết kế để giúp các nhà quảng cáo nhắm mục tiêu vào các khách hàng tiềm năng. John Wanamaker, một trong những người tạo ra cấu trúc quảng cáo hiện đại trên các phương tiện truyền thông, đã từng nói: “Một nửa số tiền tôi chi cho quảng cáo là lãng phí; vấn đề là, tôi không biết là nửa nào.”

Đây là vấn đề cơ bản với quảng cáo. Đặt quảng cáo trên trang web, bất kì ai truy cập trang web đó đều sẽ thấy quảng cáo và bạn trả tiền cho mỗi lần hiển thị. Nếu chỉ một phần nhỏ trong số họ là những khách hàng tiềm năng, thì sự sẵn sàng trả tiền cho mỗi lần hiển thị quảng cáo của bạn sẽ thấp. Đó là vấn đề của bạn với tư cách là nhà quảng cáo và với trang web đang cố gắng kiếm tiền từ quảng cáo.

Một giải pháp là tập trung vào việc xây dựng trang web thu hút người xem với những sở thích cụ thể - thể thao, tài chính... - tỷ lệ khách hàng tiềm năng sẽ cao hơn đối với một số nhà quảng cáo. Trước khi có sự bùng nổ của Internet, đây là một tính năng cốt lõi của quảng cáo, dẫn đến sự gia tăng của các tạp chí, các kênh truyền hình cáp và các chuyên mục báo cho ô tô, thời trang, bất động sản và đầu tư. Tuy nhiên, không phải mọi phương tiện truyền thông đều có thể điều chỉnh nội dung của nó theo cách này.

Thay vào đó, nhờ vào các sự đổi mới trong tìm kiếm trình duyệt web, chủ yếu là “cookie”, các nhà quảng cáo có thể theo dõi người dùng theo thời gian và

trên các trang web. Họ sau đó có khả năng nhắm mục tiêu quảng cáo tốt hơn. Các cookie lưu trữ thông tin về số lượng khách truy cập trang web nhưng quan trọng nhất là thông tin về loại trang web, bao gồm cả trang web mua sắm, họ thường xuyên vào. Nhờ có công nghệ theo dõi này, khi bạn truy cập một trang web để tìm kiếm một chiếc quần mới, bạn có thể thấy rằng một phần không cân xứng lượng quảng cáo bạn thấy, bao gồm trên các trang web hoàn toàn không liên quan, đều là về quần.

Bất kỳ trang web nào cũng có thể cài đặt cookie nhưng cookie không nhất thiết có giá trị đối với trang web đó. Thay vào đó, các trang web cung cấp cookie để bán cho những quảng cáo trao đổi (hoặc đôi khi là bán trực tiếp cho các nhà quảng cáo) để họ có thể nhắm mục tiêu quảng cáo của họ tốt hơn. Trang web bán dữ liệu về số lượt khách truy cập của họ cho các công ty đặt quảng cáo.

Các công ty mua dữ liệu vì họ không thể tự thu thập dữ liệu. Cũng không có gì đáng ngạc nhiên khi họ mua dữ liệu để xác định những khách hàng có giá trị cao. Họ cũng có thể mua dữ liệu để giúp họ tránh việc quảng cáo tới những khách hàng có giá trị thấp. Cả hai loại dữ liệu đều có giá trị giúp công ty tập trung quảng cáo của họ vào các khách hàng có giá trị.[6]

Nhiều công ty tiên phong về AI, bao gồm Google, Facebook và Microsoft, đã xây dựng hoặc mua các mạng lưới quảng cáo của riêng họ để họ có thể sở hữu loại dữ liệu có giá trị này. Họ quyết định rằng việc sở hữu dữ liệu này là xứng đáng với chi phí mua lại nó. Đối với những công ty khác, dữ liệu quảng cáo ít quan trọng hơn, vì vậy họ trao đổi quyền kiểm soát dữ liệu để tránh phát sinh chi phí cao khi thu thập nó; do vậy dữ liệu quảng cáo như vậy vẫn nằm bên ngoài những ranh giới của các công ty này.

Bán những sự dự đoán

Google, Facebook, Microsoft và một số công ty khác có dữ liệu đặc biệt hữu ích về những sở thích của người tiêu dùng trực tuyến. Thay vì chỉ bán dữ liệu, họ tiến thêm một bước nữa để đưa ra sự dự đoán cho các chuyên gia quảng cáo. Ví dụ, Google, thông qua việc tìm kiếm, YouTube và mạng lưới quảng cáo của nó, có lượng dữ liệu phong phú về nhu cầu của người dùng. Nó không bán dữ liệu. Tuy nhiên, thực tế, nó lại bán những sự dự đoán mà dữ liệu tạo ra cho các nhà quảng cáo như là một phần của dịch vụ đi kèm. Nếu

bạn quảng cáo thông qua mạng lưới của Google, quảng cáo của bạn sẽ được hiển thị cho người dùng rằng mạng lưới dự đoán khả năng cao đã bị ảnh hưởng bởi quảng cáo. Quảng cáo thông qua Facebook hoặc Microsoft mang lại những kết quả tương tự. Không có quyền truy cập trực tiếp vào dữ liệu, các nhà quảng cáo mua lại sự dự đoán.

Dữ liệu độc đáo rất quan trọng để tạo ra lợi thế chiến lược. Nếu dữ liệu không độc đáo, thật khó để xây dựng một doanh nghiệp dựa vào máy dự đoán. Không có dữ liệu, sẽ không có con đường thực sự đến học hỏi, và như vậy AI sẽ không phải là cốt lõi trong chiến lược của bạn. Như đã được nêu trong ví dụ về những mạng lưới quảng cáo, sự dự đoán vẫn có thể hữu ích. Chúng cho phép người quảng cáo nhắm tới khách hàng có giá trị cao nhất. Do đó, sự dự đoán tốt hơn có thể giúp tổ chức, ngay cả khi dữ liệu và sự dự đoán không có khả năng là nguồn lợi thế chiến lược.[7]

Cả dữ liệu và sự dự đoán nằm ngoài những ranh giới của tổ chức, nhưng nó vẫn có thể sử dụng sự dự đoán.

Ảnh hưởng chính ở đây là dữ liệu và máy dự đoán là những yếu tố bổ sung. Do đó, việc thu mua hoặc phát triển AI sẽ có giá trị hạn chế trừ khi bạn có dữ liệu để cung cấp cho nó. Nếu dữ liệu đó nằm ở các công ty khác, bạn cần một chiến lược để có được nó. Nếu dữ liệu ở chỗ một nhà cung cấp độc quyền, thì bạn có nguy cơ phải để nhà cung cấp đó đánh giá sự phù hợp trong giá trị AI của bạn. Nếu dữ liệu thuộc về đối thủ cạnh tranh, không chiến lược nào đáng để bạn bỏ tiền ra mua lại từ họ cả. Nếu dữ liệu nằm ở chỗ người tiêu dùng, họ có thể trao đổi để đổi lấy một sản phẩm hoặc một dịch vụ chất lượng tốt hơn.

Tuy nhiên, trong một số trường hợp, bạn và những người khác có thể có cùng loại dữ liệu cùng giá trị; do đó, một trao đổi dữ liệu có thể sẽ được thực hiện. Trong các tình huống khác, dữ liệu có thể thuộc về nhiều nhà cung cấp, trong trường hợp đó, bạn có thể cần sắp xếp mua một tập hợp các dữ liệu và sự dự đoán.

Việc bạn tự thu thập dữ liệu và đưa ra sự dự đoán hay mua chúng từ những người khác phụ thuộc vào tầm quan trọng của máy dự đoán đối với công ty của bạn. Nếu máy dự đoán là thông tin đầu vào mà bạn có thể loại bỏ khỏi quá trình, thì bạn có thể coi nó như loại năng lượng thông thường và mua nó từ thị trường. Ngược lại, nếu máy dự đoán là trung tâm trong chiến lược của

công ty bạn, bạn cần kiểm soát dữ liệu để cải thiện máy, vì vậy cả dữ liệu và máy dự đoán cần phải được thu thập trong nội bộ.

NHỮNG ĐIỂM CHÍNH

• Một lựa chọn chiến lược quan trọng là xác định nơi công việc kinh doanh của bạn kết thúc và nơi một công việc kinh doanh khác bắt đầu - quyết định ranh giới của công ty (ví dụ: đối tác hãng hàng không, sự thu mua các bộ phận của ô tô). Sự không chắc chắn ảnh hưởng đến sự lựa chọn này. Bởi vì máy dự đoán làm giảm sự không chắc chắn, chúng có thể ảnh hưởng đến ranh giới giữa tổ chức của bạn với những tổ chức khác.

• Bằng cách giảm thiểu sự không chắc chắn, máy dự đoán sẽ tăng khả năng soạn thảo hợp đồng, và do đó gia tăng động lực cho các công ty để ký hợp đồng cho cả nguồn thiết bị và lao động dựa vào dữ liệu, sự dự đoán và hành động. Tuy nhiên, máy dự đoán không khuyến khích các công ty ký hợp đồng lao động tập trung vào sự đánh giá. Chất lượng đánh giá rất khó để xác định trong một hợp đồng và rất khó để theo dõi. Nếu sự đánh giá được phân loại rõ ràng, nó có thể được lập trình và chúng ta sẽ không cần đến con người nữa. Vì sự đánh giá có thể sẽ đóng vai trò quan trọng trong lao động khi AI bị phân chia, việc tuyển dụng trong nội bộ sẽ gia tăng và việc thuê ngoài sẽ giảm.

• AI khuyến khích việc sở hữu dữ liệu. Tuy nhiên, việc ký kết hợp đồng để thu thập dữ liệu bên ngoài có thể cần thiết khi dữ liệu cung cấp không mang tính chiến lược thiết yếu cho tổ chức của bạn. Trong những trường hợp như vậy, tốt nhất là nên mua sự dự đoán trực tiếp thay vì mua dữ liệu và sau đó tự dự đoán.

17Chiến lược học hỏi của bạn

V

ào tháng 3 năm 2017, trong bài phát biểu quan trọng tại sự kiện I/O hằng năm, CEO Sundar Pichai của Google thông báo rằng công ty đang chuyển từ một “thế giới ưu tiên di động sang thế giới ưu tiên AI”. Đi theo sau đó là một loạt các thông báo liên quan đến AI theo nhiều cách khác nhau: từ sự phát triển của các con chip chuyên dụng để tối ưu hóa việc máy tự học, tới việc sử dụng học sâu trong những ứng dụng mới bao gồm nghiên cứu ung thư, cho đến việc đưa trợ lý được thúc đẩy bởi AI của Google áp dụng vào nhiều thiết bị nhất có thể. Pichai tuyên bố rằng công ty đã chuyển từ “tìm kiếm và sắp xếp lại thông tin của thế giới sang AI và máy tự học”.

Thông báo này mang tính chiến lược hơn là sự thay đổi tầm nhìn cơ bản. Người sáng lập Google Larry Page đã vạch ra con đường này vào năm 2002: Chúng tôi không phải lúc nào cũng sản xuất những gì mà mọi người muốn. Đó là điều mà chúng tôi luôn trăn trở. Nó thực sự rất khó. Để làm được điều đó, bạn phải thông minh, bạn phải hiểu tất cả mọi thứ trên thế giới, bạn phải hiểu những câu hỏi. Những gì chúng tôi đang cố gắng làm là tạo ra trí tuệ nhân tạo… Công cụ tìm kiếm tối ưu phải thật thông minh. Và vì vậy chúng tôi làm việc để tiến gần hơn tới điều đó.[1]

Nói theo cách này, Google đã tự nhận thấy họ đi trên con đường xây dựng trí tuệ nhân tạo trong nhiều năm. Chỉ gần đây họ mới có công khai việc áp dụng các kỹ thuật AI vào mọi việc của công ty.

Google không đơn độc trong việc thực hiện cam kết chiến lược này. Trong cùng tháng đó, Microsoft đã tuyến bố ý định của mình, bỏ qua “ưu tiên điện thoại” và “ưu tiên đám mây” sang “ưu tiên AI”.[2] Nhưng ưu tiên AI có nghĩa là gì?

Đối với cả Google và Microsoft, phần đầu tiên trong sự thay đổi của họ - không còn ưu tiên điện thoại di động nữa - cho chúng ta một đầu mối. Để ưu tiên điện thoại tức là cần thúc đẩy lưu lượng dữ liệu đến trải nghiệm điện thoại và tối ưu hóa giao diện cho người dùng ngay cả khi trang web và những

nền tảng khác bị tổn hại về chi phí. Chấp nhận sự tổn hại này là điều khiến cho nó trở thành chiến lược. “Làm tốt trên thiết bị di động” là mục tiêu nhắm tới, nhưng nếu nói rằng bạn sẽ làm như vậy ngay cả khi nó gây hại đến những kênh khác thì quả là một cam kết thực sự.

Điều này có ý nghĩa gì trong bối cảnh ưu tiên AI? Giám đốc nghiên cứu tại Google, Peter Norvig đưa ra câu trả lời:

Với việc thu thập thông tin, bất cứ điều gì trên 80% gợi nhớ và chính xác là khá tốt - không phải mọi sự gợi ý đều phải hoàn hảo, vì người dùng có thể bỏ qua những sự gợi ý không tốt. Với sự hỗ trợ, có một rào cản cao hơn xuất hiện. Bạn sẽ không sử dụng một dịch vụ đặt phòng mà sai đến 20% tổng số lần, hoặc thậm chí 2% tổng số lần. Vì vậy, một người trợ lý cần chính xác hơn nhiều, và thông minh hơn cũng như có ý thức tốt hơn về tình hình. Đó là những gì chúng tôi gọi là “ưu tiên AI”.[3]

Đó là câu trả lời hay cho một nhà khoa học máy tính. Nó nhấn mạnh cụ thể vào hiệu suất kỹ thuật và đặc biệt là độ chính xác. Nhưng mệnh đề này hàm ý một điều gì khác nữa. Nếu AI được đặt lên hàng đầu (tối đa hóa sự chính xác của sự dự đoán), thì điều gì sẽ đứng thứ hai?

Chuyên gia kinh tế biết rằng bất kỳ tuyên bố nào về việc “chúng tôi sẽ chú ý hơn vào X” đồng nghĩa với một sự đánh đổi. Vậy chúng ta cần làm gì để nhấn mạnh độ chính xác của sự dự đoán? Câu trả lời của chúng tôi đến từ khuôn mẫu kinh tế quan trọng của của chúng tôi: ưu tiên AI đồng nghĩa với việc dành nguồn lực để thu thập dữ liệu và học tập (một mục tiêu lâu dài) khi tính đến những sự cân nhắc quan trọng ngắn hạn như trải nghiệm của khách hàng ngay lập tức, doanh thu và số người dùng.

Một sự phá vỡ nhẹ

Việc áp dụng chiến lược ưu tiên AI là một sự cam kết nhằm ưu tiên chất lượng của sự dự đoán và hỗ trợ cho quá trình máy tự học, ngay cả với chi phí của những yếu tố ngắn hạn như mức độ hài lòng của khách hàng và hiệu suất hoạt động. Việc thu thập dữ liệu đồng nghĩa với việc khai thác AI có chất lượng dự đoán vẫn chưa ở mức tối ưu. Thế lưỡng nan chiến lược trọng tâm là liệu có nên ưu tiên cho việc học hay thay vào đó, bảo vệ những người khác khỏi nguy cơ phải hy sinh hiệu suất.

Những doanh nghiệp khác nhau sẽ tiếp cận tình huống nan giải này và đưa ra lựa chọn khác nhau. Nhưng tại sao Google, Microsoft và các công ty công nghệ khác lại lựa chọn ưu tiên AI? Đó có phải là điều mà các doanh nghiệp khác có thể làm theo không? Hay có điều gì đó đặc biệt về những công ty đó?

Một đặc điểm độc nhất của những công ty này là họ đã thu thập và tạo ra rất nhiều dữ liệu kỹ thuật số và hoạt động trong những môi trường không chắc chắn. Vậy nên, máy dự đoán có khả năng đem lại những công cụ mà họ sẽ sử dụng rộng rãi trong mọi sản phẩm của doanh nghiệp. Một cách nội hàm, các công cụ bao gồm sự dự đoán tốt hơn và có giá thành rẻ hơn đang rất cần thiết. Cùng với đó là lợi thế của bên cung cấp. Những công ty này đã sở hữu những nhân tài kỹ thuật mà họ có thể sử dụng để phát triển máy tự học và những ứng dụng của nó.

Những công ty này cũng giống như những người nông dân ở Iowa ở chương 15. Nhưng những công nghệ do AI thúc đẩy bộc lộ một đặc điểm quan trọng khác. Khi việc học tốn thời gian và thường dẫn đến hiệu suất kém hơn (đặc biệt với khách hàng), nó có cùng các tính năng như những gì Clay Christensen đã gọi là “công nghệ siêu đổi mới”, có nghĩa là một số công ty có uy tín sẽ thấy khó khăn trong việc áp dụng những loại công nghệ như vậy một cách nhanh chóng.[4]

Hãy xem xét một phiên bản AI mới của một sản phẩm hiện có. Để phát triển sản phẩm thì cần người dùng. Những người dùng sản phẩm AI đầu tiên sẽ có trải nghiệm tồi vì AI cần học hỏi. Một công ty có thể có một cơ sở khách hàng vững chắc, và từ đó sẽ có những khách hàng sử dụng sản phẩm đó và cung cấp dữ liệu đào tạo. Tuy nhiên, những khách hàng đó hài lòng với sản phẩm hiện tại và có thể không chấp nhận chuyển đổi sang một sản phẩm AI tạm thời kém chất lượng hơn.

Đây là “thế lưỡng nan của người đổi mới” kinh điển mà ở đó các công ty có uy tín không muốn làm gián đoạn mối quan hệ hiện có của họ với khác hàng, thậm chí ngay cả khi thu được lợi ích về lâu dài. Thế lưỡng nan của người đổi mới xảy ra bởi vì, khi họ lần đầu xuất hiện, những sự đổi mới có thể không đủ tốt để phục vụ khách hàng của các công ty uy tín trong ngành công nghiệp, nhưng họ có thể làm đủ tốt để cung cấp cho một công ty mới khởi nghiệp một lượng khách hàng vừa đủ trong một số khu vực thích hợp để xây

dựng một sản phẩm. Theo thời gian, công ty khởi nghiệp có thêm kinh nghiệm. Cuối cùng, công ty khởi nghiệp đã học đủ để tạo ra một sản phẩm đủ mạnh, lấy đi phần lớn các khách hàng của đối thủ. Đến thời điểm đó, công ty lớn hơn đã bị vượt quá xa và công ty khởi nghiệp cuối cùng sẽ thống trị.

Thế lưỡng nan của người đổi mới sẽ không còn tiến thoái lưỡng nan khi công ty phải đối mặt với sự cạnh tranh gay gắt, đặc biệt là nếu sự cạnh tranh đó đến từ những người mới tham gia, và họ cũng không phải đối mặt với những ràng buộc liên quan đến việc phải thỏa mãn lượng khách hàng hiện tại. Trong tình huống đó, cái giá của việc không làm gì là quá cao. Sự cạnh tranh như vậy gợi ý một phương trình hướng tới việc chấp nhận công nghệ siêu đổi mới nhanh chóng ngay cả khi bạn là một công ty có uy tín. Nói một cách khác, đối với những công nghệ như AI, khi mà ảnh hưởng lâu dài có thể rất lớn, một sự gián đoạn nhẹ cũng có thể dẫn đến sự áp dụng sớm, ngay cả bởi những cái tên nổi tiếng.

Việc học có thể tốn rất nhiều dữ liệu và thời gian trước khi những sự dự đoán của máy trở nên hoàn toàn chính xác. Việc máy dự đoán hoạt động tốt ngay khi xuất kho là việc rất hiếm. Ai đó bán cho bạn một phần mềm được hỗ trợ bởi AI có thể đã hết mình thực hiện công việc đào tạo. Nhưng khi bạn muốn quản lý AI vì một mục đích quan trọng cho doanh nghiệp của bạn, thực sự không có giải pháp tức thời nào cho bạn. Bạn sẽ không cần một hướng dẫn sử dụng nhiều như một hướng dẫn đào tạo. Sự đào tạo này sẽ cần tới một số cách thu thập dữ liệu và cải thiện AI.[5]

Con đường học hỏi

"Học hỏi thông qua sử dụng" là một thuật ngữ mà chuyên gia lịch sử kinh tế Nathan Rosenberg đặt ra để mô tả hiện tượng khi các công ty cải thiện thiết kế sản phẩm của họ thông qua việc tương tác với người dùng.[6] Những ứng dụng chính của ông liên quan với hiệu suất của máy bay, trong đó nhiều thiết kế bảo thủ ban đầu đã nhường chỗ cho những thiết kế tốt hơn với công suất lớn hơn và hiệu quả cao hơn khi các nhà sản xuất máy bay sửa đổi thông qua việc tương tác. Các nhà sản xuất bắt đầu sớm có lợi thế bởi vì họ học hỏi được nhiều hơn. Tất nhiên, những sự học hỏi như vậy mang lại lợi thế chiến lược trong nhiều bối cảnh khác nhau. Chúng đặc biệt quan trọng đối với máy dự đoán, mà trên thực tế là dựa vào máy tự học.

Cho đến nay, chúng tôi vẫn chưa dành nhiều thời gian phân biệt các loại hình học tập cấu thành máy tự học. Chúng tôi tập trung chủ yếu vào việc *học có giám sát*. Bạn sử dụng kỹ thuật này khi bạn đã có đủ dữ liệu về những gì bạn đang cố gắng dự đoán; ví dụ, bạn có hàng triệu hình ảnh và bạn đã biết rằng trong đó có chứa một con mèo hoặc một khối u; bạn huấn luyện AI dựa trên kiến thức đó. Việc học có giám sát là một phần quan trọng chúng tôi làm với tư cách giáo sư; chúng tôi trình bày những tài liệu mới bằng cách cho các sinh viên thấy vấn đề và giải pháp.

Ngược lại, điều gì sẽ xảy ra khi bạn không có đủ dữ liệu về những gì bạn đang cố gắng dự đoán, nhưng bạn có thể biết rằng, dựa trên thực tế, mình đã đúng bao nhiêu phần không? Trong tình huống đó, như chúng tôi đã thảo luận trong chương 2, những nhà khoa học máy tính đã khai thác các kỹ thuật học tăng cường. Trong AI, nhiều tiến bộ trong việc học tăng cường đã đến từ việc dạy máy chơi trò chơi. DeepMind đã cho AI của họ một bộ điều khiển các trò chơi điện tử như Breakout và "tặng thưởng" AI nếu nó đạt được điểm số cao hơn mà không có bất kỳ hướng dẫn nào khác. AI đã học cách chơi hàng loạt các trò chơi Atari tốt hơn những người chơi giỏi nhất. Đây là việc học thông qua sử dụng. AI đã chơi trò chơi hàng ngàn lần và học cách để chơi tốt hơn, giống như con người, ngoại trừ AI có thể chơi nhiều trò chơi hơn và nhanh hơn bất kỳ ai.[7]

Việc học xảy ra khi máy có một số bước đi nhất định và sau đó sử dụng dữ liệu cùng với trải nghiệm trong quá khứ (những nước đi và điểm số) để dự đoán nước đi nào sẽ dẫn đến sự tăng điểm số lớn nhất. Cách duy nhất để học là chơi. Không có con đường học tập, máy sẽ không thể chơi tốt hay cải thiện theo thời gian. Con đường học tập như vậy quả thực rất tốn kém.

Khi nào cần triển khai

Những người đã quen với sự phát triển phần mềm biết rằng việc mã hóa cần thử nghiệm bao quát để xác định lỗi. Trong một số trường hợp, các công ty phát hành phần mềm cho người dùng để giúp tìm ra các lỗi có thể xuất hiện khi sử dụng. Cho dù bằng cách "dùng thử" (sử dụng phiên bản đầu tiên của phần mềm trong nội bộ) hay "thử nghiệm beta" (mời những người thực hiện kiểm tra phần mềm), các hình thức học tập thông qua sử dụng này liên quan đến sự đầu tư ngắn hạn vào việc học để giúp sản phẩm cải thiện theo thời

gian. Chi phí đào tạo ngắn hạn để thu được lợi ích dài hạn này tương tự như cách con người học hỏi để thực hiện công việc của họ tốt hơn.

Phi công máy bay dân dụng cũng tiến bộ từ những kinh nghiệm làm việc. Vào ngày 15 tháng 1 năm 2009, khi chuyến bay US Airways 1549 bị cản trở bởi một đàn ngỗng Canada, khiến tất cả các động cơ bị tắt, đội trưởng Chesley "Sully" Sullenberger đã hạ cánh máy bay một cách thần kỳ trên sông Hudson, cứu mạng sống của tất cả 155 hành khách. Hầu hết các phóng viên đều cho rằng chính những kinh nghiệm của ông đã giúp thực hiện được điều này. Ông đã được ghi nhận tổng cộng 19.663 giờ bay, bao gồm việc lái 4.765 giờ chiếc Airbus A320. Bản thân Sully cũng nói: "Một cách để nhìn nhận điều này có lẽ là trong 42 năm, tôi đã thường xuyên tích lũy những khoản tiền nhỏ trong ngân hàng kinh nghiệm, học tập và đào tạo. Và vào ngày 15 tháng 1, quỹ có đủ tiền nên tôi có thể rút một khoản lớn."[8] Sully và tất cả hành khách của ông đã được hưởng lợi từ hàng nghìn người mà ông đã bay cùng trước đó.

Sự khác biệt giữa những kỹ năng của phi công trong những gì được coi là "đủ tốt để bắt đầu" dựa trên sự bỏ qua sai lầm. Rõ ràng, khả năng bỏ qua sai lầm cho phi công của chúng ta thấp hơn nhiều. Chúng ta cảm thấy an tâm hơn khi giấy chứng nhận phi công được cấp bởi Bộ Giao thông Hoa Kỳ và đòi hỏi kinh nghiệm tối thiểu 1.500 giờ bay, 500 giờ bay xuyên quốc gia, 100 giờ bay đêm và 75 giờ hoạt động bộ máy, cho dù các phi công tiếp tục học hỏi từ kinh nghiệm khi làm việc. Chúng ta có những định hướng đủ tốt khác nhau khi nói về yêu cầu đào tạo con người trong những công việc khác nhau. Điều này cũng đúng đối với máy học.

Các công ty thiết kế hệ thống để đào tạo nhân viên mới cho đến khi họ đủ tốt và bắt đầu làm việc, biết rằng họ sẽ làm tốt hơn khi họ học hỏi từ kinh nghiệm làm việc. Nhưng xác định điều gì đủ tốt là một quyết định quan trọng. Trong trường hợp của máy dự đoán, nó có thể là một quyết định chiến lược quan trọng liên quan đến thời gian: khi chuyển từ đào tạo nội bộ tới việc học hỏi từ công việc. Không có câu trả lời có sẵn nào cho điều gì cấu thành nên sự đủ tốt của máy dự đoán, chỉ có các sự đánh đổi. Thành công với máy dự đoán đòi hỏi chúng ta coi những sự đánh đổi này một cách nghiêm túc và tiếp cận chúng một cách có chiến lược.

Đầu tiên, con người có thể bỏ qua những sai sót nào? Chúng ta có khả năng cao sẽ bỏ qua sai sót của một vài máy dự đoán và khả năng thấp đối với những máy còn lại. Ví dụ, ứng dụng Inbox của Google đọc email của chúng ta, sử dụng AI để dự đoán chúng ta muốn trả lời như thế nào và tạo ra ba câu trả lời để chúng ta chọn. Nhiều người dùng thích thú sử dụng ứng dụng này cho dù nó có đến 70% khả năng thất bại (ở thời điểm viết cuốn sách này, câu trả lời do AI tạo ra chỉ hữu ích với chúng ta trong 30% tổng thời gian). Lý do cho khả năng bỏ qua sai sót cao như vậy là bởi lợi ích của việc giảm thiểu sự đánh máy và gõ vượt trội hơn so với chi phí cung cấp gợi ý và lãng phí thời gian nhìn vào màn hình khi câu trả lời được dự đoán là sai.

Ngược lại, chúng ta có khả năng không bỏ qua sai sót trong lĩnh vực lái xe tự động. Thế hệ xe tự động đầu tiên, mà chủ yếu do Google tiên phong, được đào tạo bởi những chuyên gia lái một số lượng xe hạn chế và đi hàng trăm, hàng ngàn cây số, giống như một phụ huynh giám sát một thiếu niên lái xe vậy. Những chuyên gia lái xe như vậy cung cấp một môi trường đào tạo an toàn, nhưng chúng cũng rất hạn chế. Máy chỉ có thể học trong một vài tình huống. Có những người phải mất tới nhiều triệu dặm đi trong môi trường và tình huống khác nhau trước khi họ học được cách đối phó với những sự cố hiếm gặp trên đường mà có thể dẫn đến tai nạn. Đối với xe tự lái, những con đường thực tế sẽ rất hiểm nguy và không thể tha thứ cho sai sót vì có thể gây nguy hiểm cho tính mạng của con người.

Thứ hai, thu thập dữ liệu người dùng trong thế giới thực quan trọng đến mức nào? Hiểu rằng việc đào tạo có thể tốn nhiều thời gian, Tesla tung ra những tính năng xe tự lái cho tất cả các mô hình gần đây của họ. Những tính năng này bao gồm một bộ cảm biến thu thập dữ liệu môi trường cùng với dữ liệu lái xe, được tải lên dữ liệu máy tự học của Tesla. Trong một thời gian rất ngắn, Tesla có thể nhận được dữ liệu đào tạo chỉ bằng cách quan sát người lái xe. Xe Tesla càng đi trên đường nhiều hơn thì máy của Tesla càng có thể học hỏi được nhiều hơn.

Tuy nhiên, bên cạnh việc thu thập dữ liệu một cách bị động khi con người lái xe Tesla, công ty cần dữ liệu lái xe tự động để hiểu cách hệ thống hoạt động. Để làm được như vậy, nó cần có những chiếc xe được lái tự động hoàn toàn để đánh giá hiệu suất, đồng thời phân tích khi nào người lái, với sự có mặt và tập trung cần thiết, chọn cách can thiệp. Mục tiêu cuối cùng của Tesla không

phải là sản xuất một phụ lái hay một thiếu niên lái xe dưới sự kiểm soát, mà là một chiếc xe được tự động hóa hoàn toàn. Điều đó yêu cầu đạt tới điểm con người cảm thấy thoải mái khi ở trong xe tự lái.

Ở đây ẩn giấu một sự đánh đổi phức tạp. Để làm tốt hơn, Tesla cần máy của họ học hỏi trong những tình huống thực. Nhưng đặt những chiếc xe hiện tại của họ vào những tình huống thực đồng nghĩa với việc đưa cho khách hàng một người lái xe khá non trẻ và thiếu kinh nghiệm, mặc dù có thể đã giỏi bằng hoặc hơn những người lái xe trẻ khác. Tuy nhiên, điều này còn nguy hiểm hơn cả thử nghiệm beta xem liệu Siri hay Alexa có hiểu bạn nói gì không hoặc nếu Google Inbox dự đoán chính xác phản hồi của bạn cho một email. Trong trường hợp của Siri, Alexa hoặc Google Inbox, một sai sót đồng nghĩa với chất lượng trải nghiệm người dùng thấp hơn. Trong trường hợp của xe tự động, một sai sót đồng nghĩa với việc đặt mạng sống con người vào nguy hiểm. Trải nghiệm đó có thể rất đáng sợ.[9] Xe có thể thoát khỏi đường cao tốc mà không cần thông báo hay nhấn phanh khi nhầm một đường hầm với một vật cản trở.

Những người lái xe lo lắng có thể sẽ không sử dụng những tính năng tự động và cản trở khả năng học của Tesla trong quá trình. Ngay cả khi công ty có thể thuyết phục một vài người thử nghiệm beta, liệu họ có thực sự muốn không? Tất nhiên những, người thử nghiệm beta cho lái xe tự động có thể là một người thích mạo hiểm. Trong trường hợp đó, cỗ máy của công ty sẽ ra sao?

Máy có thể học nhanh hơn với nhiều dữ liệu hơn và khi máy được sử dụng trong thực tế, nó sẽ tạo ra nhiều dữ liệu hơn. Tuy nhiên, những điều tồi tệ có thể xảy ra trong thế giới thực và làm tổn hại đến thương hiệu của công ty. Đưa những sản phẩm vào thực tế sớm hơn có thể đẩy nhanh quá trình học nhưng lại gây rủi ro tổn hại đến thương hiệu (và có lẽ là khách hàng); đưa chúng ra muộn hơn sẽ làm giảm quá trình học nhưng có đủ thời gian để cải thiện sản xuất và bảo vệ thương hiệu (và, một lần nữa, có lẽ là cả khách hàng nữa).

Với một số sản phẩm, ví dụ như Google Inbox, câu trả lời cho sự đánh đổi dường như rõ ràng vì chi phí của hiệu suất kém rất thấp và lợi nhuận từ việc học hỏi từ khách hàng sử dụng là rất cao. Việc triển khai loại hình sản phẩm này trong thế giới thực sớm hơn là hợp lý. Với những sản phẩm khác, ví dụ

như xe, câu trả lời mơ hồ hơn. Khi càng nhiều công ty khắp các ngành công nghệ khác nhau cố gắng tìm cách tận dụng lợi thế của máy tự học, những chiến lược liên quan đến việc xử lý sự đánh đổi sẽ trở nên nổi bật hơn.

Học bằng cách mô phỏng

Một bước trung gian có thể làm giảm sự đánh đổi là sử dụng những môi trường mô phỏng. Khi phi công tập luyện, trước khi họ lên máy bay thật, họ dành hàng trăm giờ ở trong những môi trường mô phỏng tinh vi và thực tế. Một cách tiếp cận tương tự cũng sẵn có cho AI. Google đào tạo AI AlphaGO của DeepMind để đánh bại những người chơi Go giỏi nhất thế giới không chỉ bằng cách quan sát hàng trăm trận đấu mà còn bằng việc đấu với phiên bản khác của chính nó.

Một hình thức tiếp cận của phương pháp này được gọi là máy tự học đối lập, AI chính được đặt mục tiêu chống lại một AI khác, cố gắng để đánh bại mục tiêu đó. Ví dụ, các chuyên gia nghiên cứu của Google cho một AI gửi tin nhắn tới một AI khác thông qua một quá trình mã hóa. Hai AI dùng chung một khóa để mã hóa và giải mã thông điệp, AI thứ ba (đối thủ) có thông điệp nhưng không có chìa khóa và cố gắng giải mã chúng. Với nhiều sự mô phỏng, đối thủ đào tạo AI chính giao tiếp bằng nhiều cách khác nhau mà khó có thể giải mã được nếu không có chìa khóa.[10]

Những tiếp cận học mô phỏng như vậy không thể xảy ra trên mặt đất; chúng đòi hỏi một thứ gì đó giống như cách tiếp cận ở một phòng thí nghiệm sản xuất ra một thuật toán máy tự học mới, có thể được sao chép và ứng dụng. Ưu điểm là máy không được đào tạo trong thực tế, vậy nên rủi ro với trải nghiệm người dùng hoặc thậm chí với bản thân người dùng, được giảm thiểu. Điểm bất lợi là những mô phỏng có thể không cung cấp dữ liệu phản hồi đủ phong phú, làm suy giảm chứ không loại bỏ nhu cầu phát hành AI sớm. Cuối cùng, bạn vẫn sẽ phải đưa AI vào thế giới thật.

Học trên mây hay học trên mặt đất

Học trong thực tế cải thiện AI. Công ty có thể sử dụng những kết quả mà máy dự đoán trải nghiệm trong thực tế để cải thiện những dự đoán sau. Thông thường, một công ty thu thập dữ liệu trong thế giới thực, những dữ liệu định nghĩa lại máy trước khi công ty đưa ra bản cập nhật mô hình dự đoán.

Autopilot của Tesla chưa từng học hỏi với người tiêu dùng thực tế. Khi nó ở ngoài thực địa, nó sẽ gửi dữ liệu về đám mây máy tính của Tesla. Tesla sau đó tổng hợp và sử dụng dữ liệu đó để nâng cấp Autopilot. Chỉ khi đó nó mới tung ra một phiên bản mới của Autopilot. Việc học diễn ra trên đám mây. Cách tiếp cận tiêu chuẩn này có lợi thế là bảo vệ người dùng khỏi những phiên bản không được đào tạo đầy đủ. Tuy nhiên, nhược điểm là AI phổ biến nằm trong các thiết bị không thể tính toán đến việc thay đổi điều kiện môi trường nhanh chóng hoặc, ít nhất, chỉ có thể làm như vậy khi dữ liệu đó được tích hợp trong thế hệ xe mới. Do đó, từ góc nhìn của người dùng, sự cải thiện này vượt bậc rõ rệt.

Ngược lại, hãy tưởng tượng nếu AI có thể học trên thiết bị và cải thiện trong môi trường đó. Nó có thể sẵn sàng phản ứng hơn với những điều kiện môi trường và tối ưu hóa bản thân cho những môi trường khác nhau. Trong những môi trường mà mọi thứ thay đổi nhanh chóng, cải thiện máy dự đoán trên chính các thiết bị của nó là rất có lợi. Ví dụ: trên những ứng dụng như Tinder (ứng dụng hẹn hò phổ biến nơi mà người dùng đưa ra sự lựa chọn bằng cách lướt trái khi không thích hoặc lướt phải khi thích), người dùng đưa ra nhiều quyết định một cách nhanh chóng. Các thuật toán dự đoán có thể tiếp nhận dữ liệu này ngay lập tức để xác định ứng viên tiềm năng nào sẽ hiển thị tiếp theo. Các sở thích của người dùng rất cụ thể và thay đổi theo thời gian, trong suốt một năm và theo thời gian trong ngày. Đến mức những người dùng trở nên gần giống nhau và có sở thích ổn định, các đám mây sẽ tiếp nhận dữ liệu và việc cập nhật sẽ hoạt động tốt hơn. Trong trường hợp một cá nhân có sở thích đặc biệt và thay đổi nhanh chóng, vậy thì khả năng điều chỉnh những dự đoán ở cấp độ của thiết bị rất hữu ích.

Các công ty cần phải đánh đổi tốc độ đưa máy dự đoán vào thế giới thực để tạo ra những dự đoán mới. Sử dụng trải nghiệm đó ngay lập tức và AI sẽ thích nghi nhanh hơn với những sự thay đổi trong điều kiện môi trường thực tế, nhưng với sự trả giá về đảm bảo chất lượng.

Quyền được học

Việc học thường đòi hỏi những khách hàng sẵn sàng cung cấp dữ liệu. Nếu chiến lược liên quan đến việc làm một cái gì đó khi phải đánh đổi một thứ khác, thì trong thế giới AI, ít có công ty nào thực hiện một cam kết mạnh mẽ và sớm hơn Apple. Tim Cook đã viết trong một phần đặc biệt nói về sự bảo

mật trên trang chủ của Apple rằng: “Ở Apple, sự tin tưởng của các bạn là tất cả đối với chúng tôi. Đó là lý do chúng tôi tôn trọng quyền riêng tư của các bạn và bảo vệ quyền riêng tư của các bạn với sự mã hóa chặt chẽ, cùng với các chính sách nghiêm ngặt, quản lý cách tất cả dữ liệu được xử lý.”[11]

Ông tiếp tục nói:

Vào một vài năm trước, người dùng Internet bắt đầu nhận ra rằng khi một dịch vụ trực tuyến là miễn phí thì bạn không phải là khách hàng. Bạn là một sản phẩm. Nhưng ở Apple, chúng tôi tin rằng một sự trải nghiệm khách hàng tuyệt vời không nên xâm phạm quyền riêng tư của các bạn.

Mô hình kinh doanh của chúng tôi rất đơn giản: Chúng tôi bán những sản phẩm tuyệt vời. Chúng tôi không xây dựng một profile dựa trên nội dung email của bạn hoặc thói quen lướt web của bạn để bán cho các nhà quảng cáo. Chúng tôi không “kiếm tiền” từ thông tin bạn lưu trữ trên iPhone hoặc trong iCloud. Và chúng tôi không đọc email hoặc tin nhắn của bạn để nhận những thông tin có thể tiếp thị tới bạn. Phần mềm và dịch vụ của chúng tôi đều được thiết kế để khiến cho thiết bị của chúng tôi tốt hơn. Chỉ đơn giản vậy thôi.[12]

Apple đã không đưa ra quyết định này vì quy định của chính phủ. Một số người cho rằng Apple đã đưa ra quyết định vì họ được cho là bị tụt lại phía sau Google và Facebook trong việc phát triển AI. Không có công ty nào, chắc chắn không phải là Apple, có thể từ chối được AI. Cam kết này sẽ khiến công việc của họ khó khăn hơn. Họ có kế hoạch thực hiện AI theo cách tôn trọng quyền riêng tư. Điều đó đánh dấu một sự cá cược chiến lược lớn mà ở đó người tiêu dùng có thể sẽ muốn kiểm soát dữ liệu của họ. Cho dù vì quyền bảo mật hay riêng tư, Apple đã đặt cược rằng cam kết của họ sẽ làm cho người tiêu dùng muốn cho phép ứng dụng AI vào thiết bị của họ nhiều hơn.[13] Apple không đơn độc trong việc đánh đối này. Salesforce, Adobe, Uber, Dropbox và nhiều công ty khác đã đầu tư rất nhiều trong vấn đề quyền riêng tư.

Sự đặt cược này là một chiến lược. Nhiều công ty khác, bao gồm Google, Facebook và Amazon, đã chọn một con đường khác bằng cách cho người dùng biết rằng họ sẽ sử dụng dữ liệu để cung cấp những sản phẩm tốt hơn.

Việc Apple tập trung vào quyền riêng tư hạn chế các sản phẩm mà họ có thể cung cấp. Ví dụ: cả Apple và Google đều có chức năng nhận diện khuôn mặt được tích hợp vào dịch vụ ảnh của họ. Để hữu ích hơn cho người tiêu dùng, các khuôn mặt phải được gắn thẻ. Google thực hiện điều này, bảo quản các thẻ gắn, ở bất kể thiết bị nào, bởi vì sự nhận diện chạy trên các máy chủ của Google. Tuy nhiên, Apple, vì lo ngại đến quyền riêng tư, đã lựa chọn nhận diện khuôn mặt ở cấp độ thiết bị. Điều đó có nghĩa là nếu bạn gắn thẻ khuôn mặt của những người bạn biết trên máy Mac của mình, các thẻ sẽ không chuyển qua iPhone hoặc iPad của bạn. Không ngạc nhiên, điều này tạo ra một tình huống mà các mối quan tâm về quyền riêng tư và khả năng sử dụng của người dùng bị hạn chế. (Vào thời điểm viết cuốn sách này, chúng tôi vẫn không biết Apple sẽ đối phó với những vấn đề này ra sao).

Chúng tôi không biết điều gì sẽ xảy ra trong thực tế. Trong mọi trường hợp, quá trình sàng lọc kinh tế của chúng tôi đã chỉ ra một cách rõ ràng rằng các sự trả giá liên quan đến việc đánh đổi quyền riêng tư của con người cho độ chính xác của máy dự đoán sẽ dẫn đến sự lựa chọn chiến lược cuối cùng. Sự tăng cường quyền riêng tư có thể cho các công ty quyền được tìm hiểu về người tiêu dùng nhưng cũng có thể có nghĩa là việc học tập không thực sự hữu ích.

Kinh nghiệm là nguồn tài nguyên khan hiếm mới

Ứng dụng điều hướng Waze thu thập dữ liệu từ những người dùng Waze khác để dự đoán vị trí của vấn đề giao thông. Nó có thể tìm ra tuyến đường nhanh nhất cho bạn. Nếu đó là tất cả những gì nó đang làm, thì sẽ không có vấn đề gì cả. Tuy nhiên, sự dự đoán thay đổi hành vi của con người, và đó là điều Waze được thiết kế để thực hiện. Khi máy nhận được thông tin từ đám đông, sự dự đoán của nó có thể bị bóp méo bởi thực tế.

Đối với Waze, vấn đề là người dùng sẽ làm theo sự hướng dẫn của nó để tránh các vấn đề giao thông, có lẽ bằng cách đi những con phố phụ. Trừ khi Waze điều chỉnh điều này, nó sẽ không bao giờ được cảnh báo rằng một vấn đề giao thông đã được giải quyết và tuyến đường bình thường lại là tuyến đường nhanh nhất. Để vượt qua trở ngại này, ứng dụng phải gửi một số người lái xe trở lại nơi tắc đường xem liệu đường đã hết tắc chưa. Điều đó thể hiện một vấn đề rõ ràng – những ai được hướng dẫn như vậy có thể là những chú cừu hy sinh cho lợi ích của đám đông. Không ngạc nhiên, điều này làm giảm

chất lượng sản phẩm của họ.

Không có cách nào dễ dàng để vượt qua sự trả giá phát sinh khi một dự đoán thay đổi hành vi của đám đông, từ đó phủ nhận AI khỏi thông tin cần thiết để hình thành sự dự đoán chính xác. Trong trường hợp này, nhu cầu của nhiều người lớn hơn nhu cầu của một vài người. Nhưng đây chắc chắn không phải là cách suy nghĩ thích hợp về việc quản lý mối quan hệ với khách hàng.

Đôi khi, để cải thiện sản phẩm, đặc biệt là khi chúng liên quan đến việc học hỏi thông qua sử dụng, điều quan trọng là thúc đẩy hệ thống để người tiêu dùng thực sự trải nghiệm điều gì đó mới mẻ mà máy có thể học hỏi. Khách hàng bị đẩy vào môi trường mới đó thường có trải nghiệm tồi tệ, nhưng những người khác sẽ được hưởng lợi từ những trải nghiệm đó. Đối với thử nghiệm beta, việc trả giá là tự nguyện, vì khách hàng chọn tham gia vào những phiên bản đầu tiên. Nhưng thử nghiệm beta có thể thu hút những khách hàng không sử dụng sản phẩm giống như cách mà những khách hàng bình thường của bạn sẽ sử dụng. Để thu thập trải nghiệm của tất cả khách hàng, đôi khi bạn cần làm giảm chất lượng sản phẩm cho khách hàng để nhận phản hồi mà sẽ có lợi cho mọi người.

Con người cũng cần có trải nghiệm

Sự khan hiếm trải nghiệm trở nên nổi bật hơn khi bạn xem xét kinh nghiệm về nguồn nhân lực của bạn. Nếu máy được trải nghiệm, thì con người có thể không. Gần đây, một số người bày tỏ lo ngại rằng tự động hóa có thể dẫn đến việc giảm bớt kỹ năng của con người.

Chuyến bay số 447 của Air France rơi xuống Đại Tây Dương trên đường từ Rio de Janeiro đến Paris năm 2009. Sự khủng hoảng bắt đầu với thời tiết xấu, nhưng lên đến đỉnh điểm khi hệ thống tự động của máy bay không hoạt động. Tại thời điểm cầm bánh lái đó, không giống như Sully trong máy bay Airways của Hoa Kỳ, theo báo cáo, một phi công tương đối thiếu kinh nghiệm đã xử lý tình hình kém. Khi một phi công giàu kinh nghiệm hơn cầm lái (trước đó anh ta đang ngủ), anh ta không thể đánh giá đúng tình hình.[14] Phi công giàu kinh nghiệm đã ngủ ít vào đêm hôm trước. Điểm mấu chốt: phi công còn non trẻ có thể đã có gần 3.000 giờ bay, nhưng đó không phải là một trải nghiệm chất lượng. Hầu hết thời gian, anh lái máy bay ở chế độ lái tự động.

Sự tự động hóa trong máy bay đã trở nên phổ biến, một phản ứng cho thực tế rằng hầu hết các tai nạn máy bay sau những năm 1970 đều là do lỗi của con người. Vì vậy, con người từ đó đã bị loại bỏ khỏi vòng điều khiển. Tuy nhiên, hậu quả mỉa mai là phi công có ít kinh nghiệm hơn và trở nên tồi tệ hơn.

Đối với chuyên gia kinh tế học Tim Harford, giải pháp là rất rõ ràng: sự tự động hóa phải được thu hẹp phạm vi. Ông lập luận rằng những gì đang được tự động hóa chỉ là những tình huống thông thường, vì vậy cần sự can thiệp của con người để biết thêm tình huống cực đoan. Nếu bạn đối phó với sự cực đoan bằng cách hài lòng với điều bình thường, thì trong đó ẩn giấu một vấn đề. Máy bay Air France đã phải đối mặt với một tình huống cực đoan mà không có sự chú ý xứng đáng từ những người có kinh nghiệm.

Harford nhấn mạnh rằng sự tự động hóa không phải lúc nào cũng dẫn đến câu hỏi hóc búa này:

Có rất nhiều tình huống trong đó sự tự động hóa không tạo ra nghịch lý như vậy. Trang web dịch vụ khách hàng có thể xử lý các khiếu nại và yêu cầu thường xuyên để nhân viên không phải thực hiện những công việc lặp đi lặp lại này mà có thể làm những công việc tốt hơn cho khách hàng, với những câu hỏi phức tạp hơn. Với một chiếc máy bay thì không như vậy. Chế độ máy tự lái và những sự hỗ trợ tinh tế hơn không giải phóng phi hành đoàn khỏi việc tập trung vào những thứ thú vị. Thay vào đó, chúng giải phóng phi hành đoàn bằng việc ngủ thiếp đi trên bảng điều khiển, theo nghĩa bóng và thậm chí cả nghĩa đen. Một sự cố khét tiếng đã xảy ra vào cuối năm 2009, khi hai phi công để chế độ tự động lái bay quá sân bay Minneapolis hơn 100 dặm. Lúc đó, họ đang mải nhìn vào máy tính xách tay của họ.[15]

Không ngạc nhiên, những ví dụ khác mà chúng tôi đã thảo luận trong cuốn sách này có xu hướng liên quan đến máy bay nhiều hơn là lĩnh vực phản hồi khách hàng, hay lĩnh vực xe tự lái. Chúng ta sẽ làm gì khi không lái xe hầu hết thời gian nhưng lại phải kiểm soát khi một sự kiện cực đoan xảy ra? Con cái chúng ta sẽ phải làm sao?

Các giải pháp được đưa ra liên quan đến việc đảm bảo con người đạt được và duy trì các kỹ năng, giảm số lượng tự động hóa để cung cấp thời gian cho con người học hỏi. Trên thực tế, trải nghiệm là một nguồn tài nguyên khan hiếm, bạn cần phân bổ một vài trải nghiệm cho con người để không bị mất kỹ năng.

Logic ngược lại cũng đúng. Để đào tạo các máy dự đoán, việc cho chúng học thông qua việc trải nghiệm các sự kiện thảm khốc chắc chắn sẽ đem lại giá trị. Nhưng nếu bạn đặt một con người vào vòng lặp, làm sao máy có thể học hỏi kinh nghiệm? Và do đó, một sự trả giá khác trong việc tạo ra một con đường học tập là việc chọn giữa kinh nghiệm của con người và máy móc.

Những đánh đổi này cho thấy ảnh hưởng của tuyên bố ưu tiên AI từ lãnh đạo của Google, Microsoft và những công ty khác. Các công ty sẵn sàng đầu tư vào dữ liệu để giúp máy của họ học hỏi. Việc cải thiện máy dự đoán được ưu tiên, ngay cả khi điều đó có nghĩa phải giảm chất lượng của trải nghiệm khách hàng hoặc sự đào tạo nhân viên. Chiến lược dữ liệu là chìa khóa cho chiến lược AI.

NHỮNG ĐIỂM CHÍNH

• Việc chuyển sang chiến lược ưu tiên AI đồng nghĩa với việc bỏ qua những ưu tiên hàng đầu trước đó. Hay nói cách khác, ưu tiên AI không phải là một câu đố - nó đại diện cho một sự đánh đổi thực tế. Chiến lược ưu tiên AI đặt việc tối đa hóa độ chính xác dự đoán trở thành mục tiêu trung tâm của tổ chức, ngay cả khi điều đó nghĩa là thỏa hiệp với những mục tiêu khác, ví dụ như tối đa hóa doanh thu, số người dùng và trải nghiệm người dùng

• AI có thể dẫn đến sự gián đoạn vì các công ty có tên tuổi thường có động lực áp dụng công nghệ yếu hơn so với các công ty khởi nghiệp. Các sản phẩm được kích hoạt bởi AI thường không tốt ở giai đoạn đầu vì cần thời gian để đào tạo một máy dự đoán thực hiện tốt như một thiết bị được mã hóa theo hướng dẫn của con người thay vì tự học. Tuy nhiên, khi được khai thác, AI có thể tiếp tục học hỏi và cải thiện, bỏ lại những sản phẩm không thông minh của đối thủ đằng sau. Các doanh nghiệp lớn rất thích cách tiếp cận chờ-và-xem kết quả, họ đứng ngoài và quan sát quá trình AI được áp dụng vào ngành công nghiệp của họ. Điều đó có thể đúng với một số công ty, nhưng những công ty khác cảm thấy khó bắt kịp một khi đối thủ cạnh tranh của họ bắt đầu đào tạo và khai thác các công cụ AI trước họ.

• Một quyết định chiến lược khác liên quan đến thời gian— khi nào nên đưa công cụ AI vào thực tế. Ban đầu, công cụ AI được đào tạo trong nhà, để tránh xa khách hàng.

Tuy nhiên, chúng học nhanh hơn khi được sử dụng thương mại vì chúng được tiếp xúc với những điều kiện hoạt động thực tế và thường có khối lượng dữ liệu lớn hơn. Lợi ích của việc khai thác sớm là học nhanh hơn và cái giá phải trả là rủi ro lớn hơn (rủi ro đối với thương hiệu hoặc sự an toàn của khách hàng thông qua việc cho khách hàng tiếp xúc với những AI còn non chưa được đào tạo tử tế). Trong một số trường hợp, sự đánh đổi rất rõ ràng, chẳng hạn như với Google Inbox, nơi mà lợi ích của việc học nhanh lớn hơn chi phí của hiệu suất không cao. Trong những trường hợp khác, chẳng hạn như lái xe tự động, sự trả giá mơ hồ hơn khi lợi ích từ việc ra mắt một sản phẩm thương mại sớm đánh bại cái giá của sự sai sót nếu sản phẩm được phát hành trước khi nó sẵn sàng.

18Quản lý rủi ro AI

L

atanya Sweeney, giám đốc công nghệ của Ủy ban Thương mại Liên bang Hoa Kỳ và hiện là giáo sư tại Đại học Harvard, đã rất ngạc nhiên khi một đồng nghiệp tìm kiếm tên của mình trên Google để tìm một trong những bài viết của bà và phát hiện một quảng cáo nói rằng bà đã bị bắt giam.[1] Sweeney bấm vào quảng cáo, trả một khoản phí và nhận ra điều mà bà đã biết: bà chưa từng bao giờ bị bắt giam. Bị hấp dẫn, bà đã nhập tên đồng nghiệp Adam Tanner và quảng cáo của cùng một công ty đã xuất hiện nhưng không có gợi ý bị bắt giam. Sau khi tìm kiếm thêm, bà đã phát triển giả thuyết rằng có thể những cái tên nghe giống như của người da đen đã kích hoạt những quảng cáo bắt giam. Sweeney sau đó thử nghiệm thêm điều này một cách có hệ thống hơn và nhận ra rằng nếu bạn tìm kiếm trên Google một tên nghe có vẻ liên quan đến người da đen như Lakisha hoặc Trevon, bạn có nhiều khả năng nhận được quảng cáo gợi ý bắt giam nhiều hơn 25% so với nếu bạn tìm kiếm một cái tên như Jill hay Joshua.[2]

Những thành kiến như vậy có khả năng gây tổn hại. Những người tìm kiếm có thể đang tìm kiếm thông tin để xem ai đó có phù hợp với công việc hay không. Nếu họ tìm thấy quảng cáo với tiêu đề như “Latanya Sweeney, bị bắt giam?”, những người tìm kiếm có thể đặt ra một số nghi ngờ. Đó là sự phân biệt đối xử vừa là sự phỉ báng.

Tại sao điều này lại xảy ra? Google cung cấp phần mềm cho phép các nhà quảng cáo kiểm tra và nhắm mục tiêu vào các từ khóa cụ thể. Các nhà quảng cáo có thể đã nhập những tên liên quan đến chủng tộc đặt cạnh quảng cáo, cho dù Google phủ nhận điều đó.[3] Một khả năng khác là mô hình xuất hiện bởi các thuật toán của Google, giúp quảng bá các quảng cáo có “điểm chất lượng” cao hơn (đồng nghĩa với việc chúng có thể được bấm vào). Các máy dự đoán có khả năng đóng vai trò trong đó. Ví dụ, nếu những nhà tuyển dụng tiềm năng tìm kiếm tên có nhiều khả năng bấm vào các quảng cáo bị bắt giam chỉ vì nghe có vẻ giống tên của người da đen hơn là những tên khác, thì điểm chất lượng liên quan tới việc liên kết những quảng cáo đó với các từ khóa đó sẽ tăng lên. Google không có ý định phân biệt đối xử, nhưng thuật toán của

nó có thể làm tăng những thành kiến đã tồn tại từ trước trong xã hội. Điều này là ví dụ minh chứng cho rủi ro của việc áp dụng AI.

Những rủi ro trách nhiệm

Sự xuất hiện của vấn đề phân biệt chủng tộc là một vấn đề xã hội, nhưng cũng là một vấn đề tiềm năng cho các công ty như Google. Họ có thể gây tranh cãi về những quy tắc chống phân biệt đối xử trong tuyển dụng. May mắn thay, khi một người như Sweeney nêu lên vấn đề, Google nhanh chóng phản hồi, điều tra và sửa chữa vấn đề.

Sự phân biệt đối xử có thể xuất hiện theo những cách tinh tế hơn rất nhiều. Các chuyên gia kinh tế Anja Lambrecht và Catherine Tucker, trong một nghiên cứu vào năm 2017, đã chỉ ra rằng quảng cáo trên Facebook có thể dẫn đến phân biệt đối xử về giới tính.[4] Họ đã đặt các quảng cáo về công việc trong khoa học, công nghệ, kỹ thuật và toán học (STEM) trên mạng xã hội và thấy rằng Facebook ít có khả năng hiển thị quảng cáo cho phụ nữ, không phải vì phụ nữ ít có khả năng bấm vào quảng cáo hơn hoặc vì chúng có thể ở những quốc gia có thị trường lao động phân biệt đối xử. Ngược lại, chính hoạt động của thị trường quảng cáo cho thấy sự phân biệt đối xử. Bởi vì những người phụ nữ trẻ có giá trị về mặt nhân khẩu học trên Facebook, hiển thị quảng cáo với họ sẽ tốn kém hơn.

Giáo sư, chuyên gi a kinh tế và luật sư của Trường Kinh doanh Harvard, Ben Edelman giải thích cho chúng tô i lý do tại sao vấn đề này có thể nghiêm trọng cho cả người tuyển dụng và Facebook. Trong khi nhiều người có xu hướng nghĩ sự phân biệt đối xử phát sinh từ việc đối xử khác nhau - đặt ra các tiêu chuẩn khác nhau cho nam và nữ - sự khác biệt về vị trí quảng cáo có thể dẫn đến điều mà luật sư gọi là “tác động khác biệt”. Một thủ tục trung lập phù hợp cho cả nam và nữ cuối cùng ảnh hưởng đến một số nhân viên – người mà có lý do để lo ngại sự phân biệt đối xử (“tầng lớp cần được bảo vệ” đối với các luật sư) khác với những người khác.

Một người hoặc một tổ chức có thể phải chịu trách nhiệm cho sự phân biệt đối xử, ngay cả khi là vô tình. Một tòa án phát hiện ra rằng Bộ Phòng cháy chữa cháy ở thành phố New York phân biệt đối xử với ứng viên da đen và gốc Latinh trong việc trở thành lính cứu hỏa với một kỳ thi tuyển sinh bao gồm nhiều câu hỏi nhấn mạnh việc đọc hiểu. Tòa án thấy rằng các câu hỏi

không liên quan đến hiệu quả làm việc của một nhân viên bộ phận phòng cháy chữa cháy và rằng các ứng viên da đen và gốc Latinh có năng lực tương đương với các ứng viên khác.[5] Vụ việc cuối cùng cũng được dàn xếp với khoảng 99 triệu đô la. Hiệu suất làm việc thấp hơn của người da đen và gốc Latinh trong kỳ thi đồng nghĩa với bộ phận đó phải chịu trách nhiệm, ngay cả khi sự phân biệt đối xử là không chủ ý.

Vì vậy, cho dù bạn nghĩ rằng mình đang đặt một quảng cáo trung lập trên Facebook, sự ảnh hưởng của việc phân biệt đối xử có thể xuất hiện bất cứ lúc nào. Với tư cách là một nhà tuyển dụng, bạn có thể sẽ phải chịu trách nhiệm. Một giải pháp cho Facebook là cung cấp các công cụ cho các nhà quảng cáo để ngăn chặn sự phân biệt đối xử.

Thách thức với AI là sự phân biệt đối xử không chủ ý như vậy có thể xảy ra mà không ai trong tổ chức nhận ra điều đó. Sự dự đoán được tạo ra bởi học sâu và nhiều công nghệ AI khác dường như được tạo ra từ một hộp đen. Việc xem xét một thuật toán hay một công thức đằng sau sự dự đoán và xác định nguyên nhân điều gì dẫn đến điều gì là không thể. Để xác định liệu AI có phân biệt đối xử hay không, bạn phải nhìn vào thông tin đầu ra. Liệu nam có kết quả khác so với nữ không? Liệu người gốc Latinh có kết quả khác so với những người khác không? Vậy còn người già hay người bị tàn tật thì sao? Liệu những kết quả khác nhau có hạn chế những cơ hội của họ?

Để ngăn chặn các vấn đề trách nhiệm pháp lý (và tránh phân biệt đối xử), nếu bạn nhận ra một sự phân biệt đối xử không chủ ý ở dữ liệu đầu ra của AI, bạn cần phải sửa nó. Bạn cần phải tìm ra lý do tại sao AI của bạn lại tạo ra sự dự đoán phân biệt đối xử như vậy. Nhưng nếu AI là một hộp đen, thì bạn có thể làm điều này như thế nào?

Một số người trong cộng đồng khoa học máy tính gọi đây là “khoa học thần kinh AI”.[6] Một công cụ quan trọng là đưa ra giả thuyết về những gì có thể tạo ra sự khác biệt, cung cấp cho AI những dữ liệu đầu vào khác nhau để kiểm tra giả thuyết và sau đó so sánh các dự đoán kết quả. Lambrecht và Tucker đã làm điều này khi họ phát hiện ra rằng phụ nữ thấy ít thấy quảng cáo STEM hơn bởi vì việc hiển thị quảng cáo cho nam giới ít tốn kém hơn. Vấn đề là hộp đen của AI không phải là một cái cớ để bỏ qua sự phân biệt đối xử tiềm tàng hoặc một cách để tránh sử dụng AI trong những tình huống mà ở đó sự

phân biệt đối xử có thể là một vấn đề. Nhiều bằng chứng cho thấy con người phân biệt đối xử thậm chí nhiều hơn máy móc. Triển khai AI đòi hỏi sự bổ sung đầu tư vào việc kiểm soát sự phân biệt đối xử, sau đó nghiên cứu để giảm thiểu bất kỳ sự phân biệt đối xử nào.

Thuật toán phân biệt đối xử có thể dễ dàng xuất hiện khi ở cấp độ hoạt động nhưng có thể sẽ gây ra những hệ quả chiến lược và bao quát hơn. Chiến lược liên quan đến việc chỉ đạo những người trong tổ chức của bạn cân nhắc các yếu tố mà có thể không rõ ràng. Điều này trở nên đặc biệt nổi bật với những rủi ro mang tính hệ thống, như sự phân biệt đối xử về thuật toán, có thể có tác động tiêu cực đến doanh nghiệp của bạn. Những hệ quả của việc tăng rủi ro có thể không trở nên rõ ràng cho đến khi quá muộn. Do đó, một công việc quan trọng dành cho các nhà lãnh đạo doanh nghiệp là dự đoán nhiều loại rủi ro khác nhau và đảm bảo rằng các quy trình được áp dụng để quản lý chúng.

Rủi ro chất lượng

Nếu bạn đang ở trong một doanh nghiệp hướng đến người tiêu dùng, bạn có thể sẽ mua quảng cáo và đã thấy một thước đo ROI của những quảng cáo đó. Ví dụ, tổ chức của bạn có thể đã thấy rằng việc trả tiền quảng cáo cho Google dẫn đến sự gia tăng trong việc bấm vào liên kết và thậm chí là việc mua hàng trên trang web. Đó là vì công ty bạn càng mua nhiều quảng cáo trên Google, càng nhiều lượt truy cập từ những quảng cáo đó hơn. Bây giờ, hãy thử sử dụng AI để xem xét dữ liệu đó và tạo ra sự dự đoán liệu quảng cáo Google mới có khả năng tăng lượt truy cập từ quảng cáo đó hay không; AI sẽ có khả năng sao lưu những tương quan tích cực mà bạn đã từng quan sát. Kết quả là, khi những người làm marketing muốn mua thêm quảng cáo Google, họ có một số bằng chứng ROI để sao lưu.

Tất nhiên, phải tốn một quảng cáo để gia tăng lượt truy cập. Có một khả năng là nếu không có quảng cáo, người tiêu dùng sẽ không bao giờ biết đến sản phẩm của bạn. Trong trường hợp này, bạn muốn đặt quảng cáo vì chúng tạo ra doanh thu mới. Một khả năng khác là quảng cáo rất dễ được khách hàng tiềm năng nhấp vào, nhưng trong trường hợp không có, khách hàng vẫn sẽ tìm đến bạn. Vì vậy, mặc dù quảng cáo có thể liên quan đến việc nhiều doanh thu hơn, nó có khả năng chỉ là viển vông. Cho dù không có quảng cáo, doanh thu cũng có thể tăng bất kể khi nào. Do đó, nếu bạn thực sự muốn biết liệu quảng cáo đó - và số tiền bạn chi tiêu vào nó – có tạo ra doanh thu mới hay

không, bạn cần phải xem xét tình hình kỹ lưỡng hơn.

Năm 2012, một số chuyên gia kinh tế làm việc cho eBay - Thomas Blake, ChrisNosko và Steve Tadelis - đã thuyết phục eBay giảm quảng cáo tìm kiếm ở Hoa Kỳ xuống còn 1/3 trong vòng một tháng.[7] Quảng cáo được đo lường bởi ROI bằng cách sử dụng thống kê số liệu truyền thống hơn 4.000%. Nếu ROI được đo là chính xác, việc thực hiện một tháng thử nghiệm sẽ tốn rất nhiều tiền của eBay.

Tuy nhiên, số liệu mà họ thu thập được đã chứng minh cho cách tiếp cận của họ. Quảng cáo tìm kiếm mà eBay đặt trên thực tế không có ảnh hưởng đến doanh thu. ROI của họ là âm. Người tiêu dùng eBay hiểu rằng, nếu họ không thấy quảng cáo trên Google, họ sẽ nhấp vào một kết quả tìm kiếm thông thường (hoặc không phải trả tiền) trên Google. Google sẽ xếp hạng cao cho danh sách eBay bất kể thế nào. Nhưng điều này cũng đúng với những thương hiệu như BMW và Amazon. Việc duy nhất quảng cáo dường như làm tốt đó là thu hút người dùng mới đến với eBay.

Mục đích của câu chuyện này là để chứng minh rằng AI - không phụ thuộc vào thử nghiệm thông thường mà phụ thuộc vào thử nghiệm về tương quan - có thể dễ dàng mắc vào những bẫy tương tự với bất kỳ ai sử dụng dữ liệu và sự thống kê đơn giản. Nếu bạn muốn biết liệu quảng cáo có hiệu quả hay không, hãy quan sát xem quảng cáo đó có dẫn đến doanh thu hay không. Tuy nhiên, bạn cũng cần phải biết điều gì sẽ xảy ra với doanh thu nếu bạn không chạy quảng cáo. AI được đào tạo trên dữ liệu liên quan đến rất nhiều quảng cáo và doanh thu không thấy được điều gì xảy ra với ít quảng cáo hơn. Dữ liệu đó bị thiếu. Những điều không biết là đã biết chính là điểm yếu của máy dự đoán và nó đòi hỏi sự đánh giá của con người để vượt qua. Tại thời điểm này, chỉ có loài người với sự sâu sắc mới có thể xác định nếu AI đang rơi vào cái bẫy đó.

Rủi ro bảo mật

Mặc dù phần mềm luôn chịu rủi ro bảo mật, với AI, những rủi ro đó xuất hiện thông qua khả năng thao tác dữ liệu. Ba lớp dữ liệu có tác động lên các máy dự đoán: đầu vào, đào tạo và phản hồi. Cả ba đều có rủi ro bảo mật tiềm ẩn.

Rủi ro dữ liệu đầu vào

Các máy dự đoán nạp dữ liệu đầu vào. Chúng kết hợp dữ liệu này với một mô hình để tạo ra sự dự đoán. Vậy nên, giống như câu thành ngữ về máy tính cũ – "rác vào, rác ra" – máy dự đoán sẽ không hoạt động nếu chúng có dữ liệu hoặc mô hình kém. Một tin tặc có thể khiến máy dự đoán không hoạt động bằng cách cho nó dữ liệu rác hoặc thao túng mô hình dự đoán. Một loại lỗi đó là sập nguồn, nó có vẻ xấu, nhưng ít nhất bạn biết là nó xảy ra. Khi ai đó thao túng máy dự đoán, bạn có thể không biết (trừ khi cho đến lúc quá muộn).

Tin tặc có nhiều cách để thao túng hoặc đánh lừa máy dự đoán. Những chuyên gia nghiên cứu của Đại học Washington cho biết, thuật toán mới của Google để phát hiện nội dung video có thể bị đánh lừa phân loại sai video bằng cách chèn những hình ảnh ngẫu nhiên vào những phần trong một giây.8 Ví dụ, bạn có thể đánh lừa AI phân loại sai video về một vườn thú bằng cách chèn hình ảnh những chiếc xe trong thời gian ngắn, con người không thể nhìn thấy những chiếc xe đó, nhưng máy tính thì có thể. Trong môi trường nơi mà những nhà sản xuất cần biết nội dung phù hợp với các nhà quảng cáo, điều này cho thấy một lỗ hổng nghiêm trọng.

Máy móc đang tạo ra các dự đoán được dùng cho quá trình đưa ra quyết định. Các công ty khai thác chúng trong những tình huống thực sự quan trọng: khi chúng ta mong chờ chúng có một ảnh hưởng thực sự lên các quyết định. Nếu không có sự tham gia của các quyết định như vậy, thì sao phải rắc rối đưa ra sự dự đoán ngay từ đầu? Những yếu tố xấu tinh vi trong trường hợp này có thể hiểu là bằng cách điều chỉnh một sự dự đoán, họ có thể điều chỉnh quyết định. Ví dụ, một bệnh nhân tiểu đường sử dụng AI để tối ưu hóa lượng insulin có thể lâm vào nguy kịch nếu AI có dữ liệu không chính xác về người đó và đưa ra những dự đoán đề xuất giảm lượng insulin trong khi đáng lẽ cần tăng. Chúng ta có nhiều khả năng sẽ khai thác máy dự đoán trong những tình huống khó dự đoán. Một yếu tố xấu có thể không tìm được chính xác dữ liệu cần có để thao túng một sự dự đoán.

Công nghệ AI sẽ phát triển cùng lúc với nhận diện danh tính. Nymi, một công ty khởi nghiệp mà chúng tôi làm việc cùng, phát triển một công nghệ sử dụng máy tự học để xác định các cá nhân thông qua nhịp tim của họ. Một vài công ty khác đang sử dụng quét võng mạc, khuôn mặt, hoặc nhận diện dấu vân tay. Bất kể thế nào, một chuỗi những công nghệ có thể xuất hiện, cho phép chúng ta cá nhân hóa AI và bảo vệ danh tính.

Trong khi các sự dự đoán cá nhân hóa có thể dễ bị thao túng bởi cá nhân, những dự đoán cá nhân có thể phải đối mặt với những rủi ro liên quan đến sự thao túng ở cấp độ dân số. Những nhà sinh thái học đã chỉ cho chúng ta rằng những quần thể đồng nhất có nguy cơ mắc bệnh cao hơn và dễ bị hủy diệt hơn.[9] Một ví dụ kinh điển là trong nông nghiệp. Nếu tất cả các nông dân trong cùng một vùng hoặc một quốc gia cùng trồng một loại giống trong cùng một mùa vụ cụ thể, họ có thể làm tốt trong thời gian ngắn. Bằng cách áp dụng giống tốt nhất, họ giảm thiểu rủi ro cá nhân. Tuy nhiên, sự đồng nhất này đã mở ra một cơ hội cho dịch bệnh hoặc thậm chí những điều kiện khí hậu bất lợi. Nếu tất cả các nông dân trồng cùng một giống, thì chúng đều dễ bị mắc cùng loại bệnh. Sự độc canh như vậy có thể có lợi cá nhân nhưng lại gia tăng rủi ro thất bại hệ thống.

Ý tưởng này cũng áp dụng cho công nghệ thông tin nói chung và máy dự đoán nói riêng. Nếu một hệ thống máy dự đoán chứng minh được bản thân nó hữu ích, thì bạn có thể áp dụng hệ thống đó ở mọi nơi trong tổ chức của bạn hoặc thậm chí thế giới. Tất cả những chiếc xe ô tô có thể áp dụng bất kỳ máy dự đoán nào có vẻ an toàn nhất. Điều này làm giảm rủi ro ở cấp độ cá nhân và gia tăng sự an toàn; tuy nhiên, nó cũng mở rộng cơ hội cho sự thất bại lớn, mặc dù cố ý hay không. Nếu tất cả các xe ô tô đều có cùng một thuật toán dự đoán, kẻ tấn công có thể khai thác thuật toán đó, thao túng dữ liệu hoặc mô hình bằng cách nào đó và khiến tất cả các xe không hoạt động cùng một lúc.

Một giải pháp dường như đơn giản cho vấn đề thất bại hệ thống này là đa dạng hóa máy dự đoán mà bạn khai thác. Điều này làm giảm rủi ro an ninh, nhưng sẽ phải giảm hiệu suất. Nó có thể làm tăng nguy cơ của những thất bại nhỏ ngẫu nhiên vì sự thiếu tiêu chuẩn hóa. Cũng như trong đa dạng sinh học, sự đa dạng của máy dự đoán liên quan đến sự đánh đổi giữa kết quả ở cấp độ cá nhân và hệ thống. Rất nhiều kịch bản cho việc thất bại hệ thống bao gồm một cuộc tấn công hàng loạt vào các máy dự đoán. Ví dụ, một cuộc tấn công vào tất cả các chiếc xe tự động đồng nghĩa với nguy cơ cho sự an toàn quốc gia.

Một cách khác để chống lại sự tấn công hàng loạt quy mô lớn, là ràng buộc thiết bị từ đám mây.[10] Chúng ta đã thảo luận về những lợi ích của việc bổ sung sự dự đoán trên mặt đất thay vì trên đám mây với mục đích đẩy mạnh việc học theo hoàn cảnh (với sự dự đoán chính xác hơn) và để bảo vệ sự riêng

tư của người tiêu dùng. Sự dự đoán trên mặt đất có một lợi ích khác. Nếu thiết bị không kết nối với đám mây, một cuộc tấn công hàng loạt sẽ trở nên khó khăn.[11] Mặc dù việc đào tạo máy dự đoán có thể xảy ra trên đám mây hoặc ở đâu đó, một khi máy được đào tạo, nó có thể dự đoán trực tiếp trên thiết bị mà không cần gửi thông tin trở về đám mây.

Rủi ro dữ liệu đào tạo

Một rủi ro khác là một ai đó có thể thẩm vấn máy dự đoán của bạn. Những đối thủ cạnh tranh của bạn có thể đảo ngược kỹ thuật thuật toán của bạn, hoặc ít nhất khiến máy dự đoán của họ sử dụng dữ liệu đầu ra từ thuật toán của bạn như là dữ liệu đào tạo. Có lẽ ví dụ nổi tiếng nhất liên quan đến phát hiện của nhóm chống spam từ Google. Máy dự đoán của Google được thiết lập kết quả giả cho rất nhiều kết quả tìm kiếm kì dị mà không tồn tại như "hiybbprqag". Vài tuần sau, nhóm đã truy vấn công cụ tìm kiếm của Bing của Microsoft. Không có gì ngạc nhiên, những kết quả tìm kiếm giả của Google cho những tìm kiếm như ""hiybbprqag" xuất hiện trong kết quả của Bing. Nhóm của Google đã phát hiện ra rằng Microsoft sử dụng thanh công cụ của mình để sao chép công cụ tìm kiếm của Google.[12]

Ở thời điểm đó, đã có rất nhiều cuộc thảo luận về liệu Microsoft làm như vậy có chấp nhận được hay không.[13] Microsoft sử dụng thanh công cụ của Google để học và phát triển những thuật toán tốt hơn cho công cụ tìm kiếm Bing của họ. Đa phần những gì người dùng đã làm là tìm kiếm trên Google và rồi nhấp vào các kết quả đó. Vậy nên khi một cụm từ tìm kiếm hiếm có và chỉ được tìm thấy trên Google (như "hiybbprqag") và nếu nó được sử dụng đủ (chính xác là điều mà những gì kỹ sư Google đang làm), máy của Microsoft sẽ học. Thật thú vị, những gì mà Microsoft đã không làm – điều mà rõ ràng có thể làm – là học cách bắt chước công cụ tìm kiếm của Google biến đổi những cụm từ vô nghĩa thành có nghĩa.[14]

Vấn đề chiến lược là khi bạn có AI (như công cụ tìm kiếm của Google), thì nếu đối thủ cạnh tranh có thể quan sát dữ liệu được nhập vào (ví dụ một truy vấn tìm kiếm) và dữ liệu đầu ra được báo cáo (ví dụ như danh sách các trang web), thì nó có nguồn thông tin thô để AI học dưới sự giám sát và xây dựng lại thuật toán. Công cụ tìm kiếm của Google có thể rất khó thực hiện, nhưng trên lý thuyết, hoàn toàn có thể.

Năm 2016, các nhà nghiên cứu khoa học máy tính đã chỉ ra rằng những thuật toán học sâu cụ thể đều dễ bị bắt chước.[15] Họ đã thử nghiệm khả năng này trên một số nền tảng máy tự học quan trọng (bao gồm máy tự học Amazon) và chứng minh rằng với số lượng truy vấn nhỏ (650 – 4.000), họ có thể đảo ngược những mô hình đó tới mức xấp xỉ gần đúng, đôi khi là hoàn hảo. Việc khai thác nhiều thuật toán máy tự học dẫn đến lỗ hổng này.

Sự bắt chước có thể dễ. Sau khi bạn đã thực hiện tất cả các công việc đào tạo AI, cả thế giới sẽ biết cách vận hành của nó và nó có thể sẽ bị sao chép. Nhưng đáng lo ngại hơn là sự chiếm đoạt kiến thức này có thể dẫn đến tình huống những yếu tố xấu thao túng sự dự đoán và quá trình học một cách dễ dàng. Một khi kẻ tấn công hiểu máy, máy trở nên dễ bị tổn thương hơn.

Về mặt tích cực, những sự tấn công đó sẽ để lại dấu vết. Việc truy vấn máy dự đoán nhiều lần để hiểu nó là cần thiết. Số lượng truy vấn bất thường hoặc sự đa dạng truy vấn bất thường sẽ đưa ra cảnh báo. Một khi cảnh báo thì việc bảo vệ máy dự đoán trở nên dễ hơn, cho dù không hẳn dễ dàng. Nhưng ít nhất bạn biết rằng một sự tấn công đang xảy đến và kẻ tấn công đã thu được gì. Sau đó, bạn có thể bảo vệ máy bằng cách chặn kẻ tấn công hoặc chuẩn bị một kế hoạch dự phòng nếu có gì đó xảy ra.

Rủi ro dữ liệu phản hồi

Máy dự đoán của bạn sẽ tương tác với những nhân tố khác (con người hoặc máy) bên ngoài doanh nghiệp của bạn, điều này tạo ra một nguy cơ khác: những nhân tố xấu có thể cung cấp dữ liệu ảnh hưởng đến quá trình học của AI. Điều này còn hơn cả thao túng một sự dự đoán, mà thay vào đó liên quan đến việc dạy máy cách dự đoán sai có hệ thống.

Một ví dụ công khai và đầy ấn tượng xảy ra vào tháng 3 năm 2016 khi Microsoft phát hành một chatbot trên Twitter tên Tay. Ý tưởng của Microsoft khá rõ ràng: để Tay tương tác với mọi người trên Twitter và quyết định xem phản hồi thế nào là tốt nhất. Ý định của nó là học cụ thể về “cuộc trò chuyện bình thường và vui vẻ”.[16] Trên giấy tờ, đây là một cách hợp lý để AI tiếp xúc với trải nghiệm nó cần để học nhanh. Tay ban đầu không hơn một chú vẹt, nhưng với mục tiêu tham vọng hơn.

Internet, tuy nhiên, không phải là một bối cảnh nhẹ nhàng. Ngay sau khi được

khởi động, mọi người bắt đầu thử thách giới hạn của Tay. Baron Memilton hỏi: “Bạn có ủng hộ nạn diệt chủng không” và Tay trả lời lại: “@Baron_von_Derp Thực sự tôi có (ủng hộ)”. Tay nhanh chóng trở thành một người phân biệt chủng tộc, phân biệt phụ nữ, ủng hộ Đức Quốc xã. Microsoft đã loại bỏ cuộc thí nghiệm.[17]

Những sự ảnh hưởng rất rõ ràng. Những đối thủ cạnh tranh hoặc những người phỉ báng có thể sẽ cố đào tạo máy dự đoán của bạn đưa ra những dự đoán tồi. Cũng giống như Tay, dữ liệu đào tạo máy dự đoán và máy dự đoán được đào tạo trong thực tế có thể sẽ gặp phải những người sử dụng máy một cách chiến lược, xấu tính hoặc không trung thực.

Đối mặt với rủi ro

Máy dự đoán mang theo rủi ro. Bất kỳ công ty nào đầu tư vào Ai sẽ đối mặt với những rủi ro này và loại bỏ tất cả chúng là không thể. Không có giải pháp dễ dàng nào nhưng ít nhất bây giờ bạn có kiến thức để dự đoán những dự đoán này. Hãy nhớ rằng những dự đoán của bạn sẽ thay đổi tùy theo nhóm người. Hãy cẩn thận những yếu tố xấu có thể truy vấn máy dự đoán của bạn để bắt chước hoặc thậm chí phá hủy chúng.

NHỮNG ĐIỂM CHÍNH

AI mang theo nhiều loại rủi ro. Chúng tôi khái quát sáu trong số những loại nổi bật nhất ở đây.

1. Những dự đoán của AI có thể dẫn đến sự phân biệt đối xử. Ngay cả khi sự phân biệt đó là vô tình, nó vẫn có trách nhiệm pháp lý.

2. AI không hoạt động hiệu quả khi dữ liệu ít. Điều này gây ra rủi ro về chất lượng, cụ thể là những điều chưa biết là đã biết, trong đó sự dự đoán được cung cấp với sự tự tin nhưng lại sai.

3. Dữ liệu đầu vào sai có thể đánh lừa máy dự đoán, khiến người dùng dễ bị tấn công bởi tin tặc.

4. Cũng giống như trong đa dạng sinh học, sự đa dạng của máy dự đoán liên quan đến sự cân bằng giữa kết quả ở cấp độ cá nhân và cấp độ hệ thống. Sự ít

đa dạng hơn có thể có lợi cho hiệu suất ở cấp độ cá nhân, nhưng gia tăng rủi ro thất bại lớn.

5. Máy dự đoán có thể bị thẩm vấn, khiến bạn dễ bị đánh cắp sở hữu trí tuệ và những kẻ tấn công phát hiện những yếu điểm của bạn.

6. Dữ liệu phản hồi có thể bị thao túng, khiến máy dự đoán học những hành vi mang tính phá hoại.

PHẦN 5XÃ HỘI



19Vượt xa hơn việc kinh doanh

N

hiều cuộc thảo luận phổ biến về AI liên quan đến những vấn đề xã hội thay vì kinh doanh. Nhiều người không chắc liệu AI có tốt hay không. CEO của Tesla, Elon Musk là một trong những người nổi tiếng, có kinh nghiệm thường xuyên rung hồi chuông cảnh tỉnh: “Tôi đã tiếp xúc với loại hình AI rất tân tiến, và tôi nghĩ mọi người nên thực sự quan tâm đến nó… Tôi liên tục rung những hồi chuông cảnh tỉnh, nhưng con người vẫn không biết phải phản ứng thế nào với robot, vì nó dường như không có thật.”[1]

Một nhà nghiên cứu uyên bác khác có ý kiến trong vấn đề này là chuyên gia tâm lý học nhận được giải Nobel, Daniel Kahnman, tác giả của cuốn sách*Thinking, Fast and Slow* (tạm dịch: Nghĩ nhanh nghĩ chậm). Năm 2017, tại một hội nghị chúng tôi tổ chức ở Toronto về kinh tế học của trí tuệ nhân tạo, ông giải thích vì sao ông nghĩ AI sẽ thông minh hơn con người:

Cách đây không lâu, một tiểu thuyết gia nổi tiếng viết cho tôi rằng ông đang có kế hoạch viết một cuốn tiểu thuyết. Nội dung của cuốn tiểu thuyết là về tam giác tình yêu giữa hai con người và một robot và điều mà ông muốn biết là robot khác với con người như thế nào.

Tôi đề xuất ba điểm khác biệt chính. Điều hiển nhiên đầu tiên: robot sẽ giỏi hơn nhiều trong việc lý luận thống kê và không bị lừa mị bởi những câu chuyện cũng như những lời thuật lại. Một điểm khác là robot sẽ có trí thông minh cảm xúc cao hơn. Điều thứ ba là robot sẽ thông minh hơn. Trí thông minh đồng nghĩa với việc không có tầm nhìn hạn hẹp. Đó là bản chất của trí thông minh; một tầm nhìn rộng. Tôi cho rằng khi nó đã học đủ, nó sẽ thông minh hơn con người bởi vì chúng ta không có tầm nhìn rộng. Chúng ta là những người suy nghĩ hạn hẹp, chúng ta là những người suy nghĩ ồn ào, và nó rất dễ tiến bộ hơn chúng ta. Tôi nghĩ là những gì chúng ta có thể làm, robot rồi cũng sẽ [học] được thôi.

Elon Musk và Daniel Kahneman đều rất tự tin về tiềm năng của AI, đồng thời cũng rất lo lắng về những tác động của việc mang nó đến với thế giới. Thiếu

kiên nhẫn về tốc độ chính phủ phản ứng lại với những tiến bộ công nghệ, các nhà lãnh đạo ngành đã đưa ra các đề xuất chính sách và, trong một vài trường hợp, đã hành động. Bill Gates ủng hộ việc đánh thuế robot thay thế lao động của con người. Bên cạnh những việc thường được coi là quyền hạn của chính phủ, những công ty khởi nghiệp cao cấp thúc đẩy Y Combinator đang thử nghiệm về việc cung cấp một khoản thu nhập cơ bản cho tất cả mọi người trong xã hội.[2] Elon Musk sắp xếp một nhóm những nhà khởi nghiệp và lãnh đạo ngành để cung cấp tài chính cho Open AI với 1 tỷ đô la để đảm bảo rằng không có công ty tư nhân nào có thể độc quyền hóa lĩnh vực này.

Những đề xuất và hành động như vậy nhấn mạnh sự phức tạp của những vấn đề xã hội này. Khi chúng ta trèo lên đỉnh của kim tự tháp, các lựa chọn trở nên phức tạp hơn một cách đáng kể. Khi nghĩ về xã hội nói chung, những vấn đề kinh tế của AI không còn đơn giản nữa.

Đây liệu có phải là dấu chấm hết cho các công việc không?

Nếu Einstein có một hiện thân hiện đại, thì đó là Stephen Hawking. Nhờ có những đóng góp to lớn của ông cho khoa học và những cuốn sách như A Brief History of Time (tạm dịch: Lược sử về thời gian), cho dù phải đấu tranh với bệnh ALS, Hawking vẫn được coi là thiên tài lỗi lạc của thế giới. Vì vậy, mọi người không ngạc nhiên khi vào tháng 12 năm 2016, ông viết: “Sự tự động hóa của các nhà máy đã làm mất đi 1/10 số công việc sản xuất truyền thống, và sự phát triển của trí tuệ nhân tạo có khả năng mở rộng sự phá hủy công việc sâu sắc đến những tầng lớp trung lưu, và chỉ những vai trò chăm sóc, sáng tạo hoặc giám sát sẽ còn sót lại.”[3]

Nhiều nghiên cứu đã tính đến sự phá hủy công việc tiềm năng do sự tự động hóa gây ra, và lần này nó không chỉ là lao động thể chất nữa mà còn là những chức năng nhận thức mà trước kia được tin rằng là miễn dịch với những tác động đó.[4] Xét cho cùng, ngựa thua cuộc vì mã lực, chứ không phải vì trí lực. Với tư cách là những chuyên gia kinh tế, chúng tôi đã nghe những lời tuyên bố này từ trước. Nhưng cho dù bóng ma của sự thất nghiệp gây ra bởi công nghiệp đã lờ mờ xuất hiện từ khi Luddites phá hủy những khung dệt từ những thế kỷ trước, tỉ lệ thất nghiệp vẫn đang ở mức thấp đáng kể.

Nhưng liệu lần này có khác? Mối lo ngại của Hawking, cùng với nhiều

người, là lần này có thể sẽ khác vì AI có thể lấy đi những lợi thế còn lại của con người.[5] Một chuyên gia kinh tế tiếp cận câu hỏi này như thế nào? Hãy tưởng tượng có một hòn đảo mới chỉ có toàn robot – Robotlandia – đột nhiên xuất hiện. Liệu chúng ta có muốn giao dịch với hòn đảo của máy dự đoán này không? Từ quan điểm giao dịch tự do, đây dường như là một cơ hội tốt. Robot tự làm tất cả công việc, giải phóng con người khỏi việc làm những gì họ giỏi nhất.

Tất nhiên, không có Robotlandia nào thực sự tồn tại, nhưng khi chúng ta có những thay đổi kỹ thuật khiến cho phần mềm có khả năng làm những công việc mới với giá thành rẻ hơn, các chuyên gia kinh tế nhìn nhận nó tương tự với việc mở cửa giao dịch với một hòn đảo tưởng tượng. Hay nói cách khác, nếu bạn ủng hộ giao dịch tự do giữa các nước, thì bạn sẽ ủng hộ giao dịch tự do với Robotlandia. Bạn ủng hộ việc phát triển AI, ngay cả khi nó thay thế một vài công việc. Nhiều thập kỷ nghiên cứu về ảnh hưởng của việc giao dịch cho thấy những công việc khác sẽ biến mất, và nhìn chung tỉ lệ có việc làm sẽ không giảm.

Mô hình của sự quyết định gợi ý những công việc mới này sẽ xuất hiện từ đâu. Con người và AI có thể làm việc cùng nhau, con người sẽ cung cấp những bổ sung cho sự dự đoán, cụ thể là, dữ liệu, sự đánh giá hoặc hành động. Ví dụ, khi giá thành của sự dự đoán trở nên rẻ hơn, giá trị của sự đánh giá gia tăng. Nhờ đó chúng ta dự đoán sự tăng trưởng của số lượng công việc liên quan đến REF.

Những công việc liên quan đến sự đánh giá khác sẽ trở nên phổ biến hơn, nhưng có lẽ đòi hỏi ít kỹ năng hơn những công việc mà AI thay thế. Nhiều công việc được trả lương cao hiện nay đòi hỏi sự đánh giá như là một kỹ năng chính, bao gồm những công việc như bác sĩ, chuyên gia phân tích tài chính và luật sư. Giống như những chỉ dẫn của máy dự đoán dẫn đến việc giảm thu nhập cho những tài xế taxi nhưng lại tăng số lượng những tài xế Uber bị trả lương thấp, chúng tôi hy vọng sẽ thấy hiện tượng tương tự như vậy trong y học và tài chính. Khi việc dự đoán được tự động hóa, nhiều người sẽ làm những công việc này, tập trung cụ thể hơn vào những kỹ năng liên quan đến đánh giá. Khi sự dự đoán không còn là một giá trị ràng buộc, nhu cầu cho những kỹ năng bổ sung phổ biến hơn có thể tăng, dẫn đến nhiều việc làm hơn nhưng với mức lương thấp hơn.

AI và con người có một điểm khác biệt quan trọng: quy mô phần mềm. Điều này có nghĩa là một khi AI giỏi hơn con người ở một công việc cụ thể, sự mất công việc sẽ xảy ra nhanh hơn. Chúng ta có thể tự tin rằng những công việc mới sẽ xuất hiện trong một vài năm nữa và con người sẽ có việc để làm, nhưng những người tìm kiếm công việc và chờ những công việc mới xuất hiện sẽ cảm thấy ít thoải mái.

Liệu sự bất bình đẳng có trở nên tồi tệ hơn không?

Công việc là một chuyện. Thu nhập có được từ công việc là chuyện khác. Việc mở cửa giao dịch thường xuyên tạo ra sự cạnh tranh và sự cạnh tranh khiến giá thành giảm. Nếu cạnh tranh với sức lao động con người thì mức lương sẽ giảm. Với giao dịch giữa các nước, nước thắng cuộc và thua cuộc từ cuộc giao dịch với máy sẽ xuất hiện: Công việc sẽ vẫn tồn tại, nhưng một vài người sẽ có những công việc ít hấp dẫn hơn bây giờ. Hay nói cách khác, nếu bạn hiểu những lợi ích của việc giao dịch tự do, thì bạn nên trân trọng những gì nhận được từ máy dự đoán. Câu hỏi chính sách quan trọng không phải là liệu AI sẽ mang lại lợi ích hay không mà là những lợi ích này sẽ được phân bổ như thế nào.

Bởi vì công cụ AI có thể được sử dụng để thay thế những kỹ năng “cao” – như trí tuệ - nên nhiều người lo lắng rằng ngay cả khi công việc tồn tại, mức lương sẽ không cao. Ví dụ, khi là Chủ tịch Hội đồng tư vấn kinh tế của Obama, Jason Furman thể hiện sự lo lắng của mình:

Mối lo lắng của tôi không phải là lần này có thể sẽ khác với AI, mà lần này có thể vẫn sẽ giống như những gì chúng ta đã trải qua trong vài thập niên vừa qua. Tranh cãi truyền thống rằng chúng ta không cần phải lo lắng về việc robot sẽ lấy đi công việc vẫn khiến chúng ta phải lo lắng rằng lý do duy nhất chúng ta vẫn sẽ còn công việc là bởi chúng ta sẵn sàng làm với mức lương thấp.[6]

Nếu phần công việc của máy tiếp tục tăng, thì mức lương của người làm sẽ giảm, trong khi những người sở hữu AI sẽ tăng. Trong cuốn sách bán chạy của mình, Capital in the Twenty-First Century (tạm dịch:Vốn của thế kỉ 21), Thomas Piketty nhấn mạnh rằng trong vài thập niên vừa qua, phần thu nhập của người lao động giảm (ở Hoa KỲ và ở bất cứ đâu) để thúc đẩy sự gia tăng của vốn. Xu hướng này rất đáng lo ngại vì nó dẫn đến sự bất bình đẳng gia

tăng. Câu hỏi quan trọng ở đây là liệu AI sẽ củng cố xu hướng này hay làm giảm thiểu nó. Nếu AI là một dạng vốn mới, hiệu quả, thì phần vốn của nền kinh tế có thể sẽ tiếp tục tăng lên với chi phí của sự lao động.

Không có giải pháp dễ dàng nào cho vấn đề này. Ví dụ, gợi ý của Bill Gates về thuế robot sẽ giảm thiểu sự bất bình đẳng nhưng sẽ khiến việc mua robot ít có lợi nhuận hơn. Vì vậy các công ty sẽ đầu tư ít hơn vào robot, năng suất sẽ chậm hơn, và chúng ta nhìn chung sẽ nghèo hơn. Chính sách giao dịch rất rõ ràng: chúng ta có chính sách có thể giảm thiểu sự bất bình đẳng nhưng kéo theo mức lương chung thấp hơn.

Xu hướng thứ hai dẫn đến việc gia tăng sự bất bình đẳng là công nghệ thường thiên vị kỹ năng. Nó làm tăng mức lương của những người có học thức một cách không cân xứng và thậm chí có thể làm giảm mức lương của những người ít có học thức hơn. Như chuyên gia kinh tế Claudia Goldin và Lawrence Katz nói, “những người với học thức và những khả năng bẩm sinh cao hơn sẽ nắm bắt được những công cụ mới và phức tạp.”[7]

Chúng ta không có lý do gì để hy vọng AI sẽ khác biệt. Những người có học thức cao có xu hướng giỏi hơn trong việc học những kỹ năng mới. Nếu những kỹ năng cần thiết để sử dụng AI hiệu quả thay đổi thường xuyên hơn, thì những người có học thức sẽ được hưởng lợi.

Chúng tôi nhận ra rằng để sử dụng AI một cách hiệu quả đòi hỏi những kỹ năng bổ sung. Ví dụ, REF phải hiểu những mục tiêu của tổ chức và khả năng của máy. Bởi vì máy có quy mô hiệu quả, nếu kỹ năng này khan hiếm, thì những kỹ sư giỏi nhất sẽ gặt hái được những lợi ích từ công việc của họ trên hàng triệu hoặc hàng tỷ máy.

Chính vì những kỹ năng liên quan đến AI hiện tại đang khan hiếm, quá trình học cho cả con người và doanh nghiệp sẽ rất tốn kém. Phần lớn lực lượng lao động được đào tạo qua nhiều thập kỷ, đồng nghĩa với việc cần phải đào tạo lại và tái cấu trúc kỹ năng. Hệ thống đào tạo công nghiệp của chúng ta không được thiết kế cho việc đó. Doanh nghiệp không nên hy vọng vào việc hệ thống thay đổi đủ nhanh để cung cấp cho họ những công nhân mà họ cần để cạnh tranh trong thời đại AI. Những thách thức về chính sách không hề đơn giản: việc nâng cao giáo dục rất tốn kém. Các chi phí này cần được thanh toán bằng cách tăng thuế hoặc bởi chính các doanh nghiệp và cá nhân. Ngay

cả khi chi phí có thể dễ dàng được xử lý, nhiều người trung niên có thể sẽ không muốn trở lại trường học. Những người bị tổn thương nhất bởi công nghệ thiên vị kỹ năng có thể sẽ là những người không sẵn sàng cho việc học tập cả đời.

Một vài công ty lớn sẽ kiểm soát mọi thứ?

Không chỉ có các cá nhân lo lắng về AI. Nhiều công ty đang lo sợ rằng họ sẽ tụt lại phía sau đối thủ trong việc đảm bảo và sử dụng AI, ít nhất một phần là do quy mô kinh tế có thể gắn liền với AI. Nhiều khách hàng đồng nghĩa với nhiều dữ liệu hơn, nhiều dữ liệu hơn đồng nghĩa với sự dự đoán AI chính xác, nhiều sự dự đoán chính xác đồng nghĩa với nhiều khách hàng hơn, và vòng tròn này tiếp tục. Trong những điều kiện thích hợp, một khi có một công ty AI dẫn đầu về hiệu suất, các đối thủ cạnh tranh của họ có thể sẽ không bao giờ bắt kịp được. Trong thí nghiệm về dự đoán chuyển hàng của Amazon trong chương 2, quy mô và lợi thế đi trước của Amazon có thể tạo ra sự dẫn đầu về độ dự đoán chính xác mà các đối thủ sẽ thấy bất khả thi để bắt kịp.

Đây không phải là lần đầu tiên mà một công nghệ mới làm gia tăng khả năng cung cấp cho các công ty lớn. AT&T kiểm soát viễn thông ở Hoa Kỳ trong hơn 50 năm. Microsoft và Intel độc quyền trong công nghệ thông tin ở những 1990 và 2000. Gần đây, Google thống trị mảng tìm kiếm và Facebook thống trị mạng truyền thông. Những công ty lớn này tăng trưởng lớn bởi những công nghệ cốt lõi cho phép họ giảm chi phí và cải thiện chất lượng tốt hơn khi họ mở rộng. Đồng thời, các đối thủ cạnh tranh nổi lên, ngay cả khi đối mặt với quy mô kinh tế lớn; hãy hỏi Microsoft (Apple và Google), Intel (AMD và ARM), và AT&T (hầu hết mọi người). Độc quyền về mặt công nghệ chỉ là tạm thời do một quá trình mà chuyên gia kinh tế Joseph Schumpeter gọi là “cơn thịnh nộ của sự hủy diệt sáng tạo”.

Với AI, việc là một công ty lớn sẽ đem lại nhiều lợi ích vì quy mô kinh tế. Tuy nhiên, nó không đồng nghĩa với việc một công ty sẽ thống trị hoặc thậm chí nếu có một công ty thống trị, điều đó sẽ không kéo dài. Trên quy mô toàn cầu, điều đó thậm chí còn đúng hơn. Nếu AI có quy mô kinh tế, điều đó sẽ không ảnh hưởng đến tất cả các ngành công nghiệp. Nếu công ty của bạn thành công và có danh tiếng, sự dự đoán chính xác có thể không phải là điều duy nhất khiến nó thành công. Khả năng hoặc tài sản có giá trị ở ngày nay có thể sẽ vẫn có giá trị khi kết hợp với AI. Đối với các công ty công nghệ mà

toàn bộ doanh nghiệp có thể dựa vào AI, quy mô kinh tế có thể dẫn đến một vài công ty chiếm ưu thế. Nhưng khi chúng ta nói đến quy mô kinh tế, chúng ta đang nói về mức quy mô thế nào?

Không có câu trả lời đơn giản nào cho câu hỏi đó, và chắc chắn chúng tôi không có dự đoán chính xác liên quan đến AI. Nhưng các chuyên gia kinh tế đã nghiên cứu quy mô kinh tế của yếu tố bổ sung quan trọng với AI: dữ liệu. Trong khi nhiều lý do có thể giải thích 70% thị trường độc chiếm của Google trong việc tìm kiếm ở Mỹ và 90% ở Liên minh châu Âu, một lời giải thích hàng đầu là Google có nhiều dữ liệu đào tạo công cụ tìm kiếm AI của hơn đối thủ của họ. Google đã thu nhập dữ liệu trong nhiều năm. Hơn nữa, phần trăm thị trường ưu thế của họ tạo ra một vòng tròn lặp lại cho quy mô dữ liệu mà những công ty khác có thể không bao giờ sánh bằng.

Hai chuyên gia kinh tế - Lesley Chiou và Catherine Tucker – nghiên cứu về các công cụ tìm kiếm tận dụng sự khác biệt trong việc lưu giữ dữ liệu.[8] Trả lời cho những đề xuất của EU năm 2008, Yahoo và Bing giảm số lượng dữ liệu mà họ lưu giữ. Google không thay đổi chính sách của họ. Những sự thay đổi này đủ để Chiou và Tucker đánh giá hiệu quả của quy mô dữ liệu trong độ chính xác tìm kiếm. Thật thú vị, họ nhận ra rằng quy mô không quan trọng lắm. So với khối lượng dữ liệu khổng lồ mà tất cả các đối thủ cạnh tranh lớn sử dụng, ít dữ liệu không có ảnh hưởng tiêu cực đến kết quả tìm kiếm. Điều này cho thấy dữ liệu lịch sử có thể ít hữu ích hơn nhiều người giả định, có thể bởi vì thế giới thay đổi quá nhanh.

Tuy nhiên, chúng tôi sẽ tiết lộ cho bạn một thông báo quan trọng. Có tới 20% tìm kiếm của Google mỗi ngày được cho là độc đáo.[9] Bởi vậy, Google có lợi thế trong việc tìm kiếm những cụm từ hiếm, nhưng trong thị trường cạnh tranh cao trong việc tìm kiếm hiện nay, một lợi thế nhỏ cũng có thể chuyển đổi thành phần trăm trên thị trường lớn hơn.

Bất chấp những đối thủ cạnh tranh tiềm năng, các công ty AI lớn sẽ vẫn bành chướng. Họ có thể mua lại các công ty khởi nghiệp trước khi chúng trở thành mối đe dọa, kiềm chế những ý tưởng mới và làm giảm năng suất về mặt lâu dài. Họ có thể đặt giá cho AI quá cao, làm ảnh hưởng đến người tiêu dùng và các doanh nghiệp khác. Việc phá vỡ độc quyền làm giảm quy mô, nhưng quy mô khiến AI tốt hơn. Một lần nữa, đưa ra chính sách không hề đơn giản.[10]

Một số quốc gia sẽ có lợi thế hơn?

Vào ngày 1 tháng 9 năm 2017, Tổng thống Nga Vladimir Putin đã khẳng định tầm quan trọng của lãnh đạo AI: "Trí tuệ nhân tạo là tương lai, không chỉ cho Nga mà còn cho cả nhân loại... Nó mang đến những cơ hội khổng lồ, nhưng cũng đem lại những mối đe dọa khó dự đoán. Ai trở thành nhà lãnh đạo trong lĩnh vực này sẽ trở thành người cai trị thế giới."[11] Liệu các quốc gia có thể hưởng lợi từ quy mô kinh tế AI như các công ty? Các quốc gia có thể thiết kế môi trường quy định cũng như chỉ đạo chính phủ chi tiêu để tăng tốc việc phát triển AI. Các chính sách có mục tiêu có thể cung cấp cho các quốc gia, và các doanh nghiệp, một lợi thế về AI.

Về phía các trường đại học và doanh nghiệp, Hoa Kỳ dẫn đầu thế giới về nghiên cứu và ứng dụng AI. Về phía chính phủ, Nhà Trắng đã công bố bốn bản báo cáo trong hai nhiệm kỳ cuối của Obama.[12] Dưới nhiệm kỳ của Obama, hầu hết các cơ quan chính phủ lớn, từ Bộ Thương mại cho đến Cơ quan An ninh Quốc gia, đã tăng tốc cho sự xuất hiện của AI ở cấp độ thương mại.

Tuy nhiên, các xu hướng đang thay đổi. Cụ thể, quốc gia lớn nhất thế giới như Trung Quốc nổi bật vì thành công trong AI. Hai công ty công nghệ định hướng AI của Trung Quốc – Tencent và Alibaba – nằm trong top 12 thế giới về giá trị và nhiều bằng chứng cho thấy sự phát triển khoa học trong AI của họ có thể sớm dẫn đầu thế giới. Ví dụ, các nghiên cứu của Trung Quốc trong hội nghị nghiên cứu AI lớn nhất tăng từ 10% năm 2012 đến 23% năm 2017. Trong khi đó, Hoa Kỳ giảm từ 41% xuống 34%.[13] Liệu tương lai của AI có "được sản xuất tại Trung Quốc"(Made in China), như tờ New York Times đề xuất không?[14] Bên cạnh việc lãnh đạo khoa học, ít nhất có thêm ba lý do chỉ ra rằng Trung Quốc đang trở thành quốc gia hàng đầu thế giới về AI.[15]

Đầu tiên, Trung Quốc chi hàng tỷ vào AI, bao gồm các dự án lớn, các công ty khởi nghiệp và nghiên cứu cơ bản. Một thành phố lớn thứ tám của Trung Quốc đã phân bổ nhiều nguồn lực vào AI hơn toàn bộ Canada. "Vào tháng 6, chính phủ của tỉnh Thiên Tân, một thành phố phía Đông gần Bắc Kinh, cho biết họ có kế hoạch thiết lập một quỹ 5 tỷ đô để hỗ trợ ngành công nghiệp AI. Họ cũng thiết lập một 'khu công nghiệp trí tuệ' hơn 20km vuông."[16] Trong khi đó, chính phủ Hoa Kỳ dường như chi tiêu ít vào khoa học hơn dưới

nhiệm kỳ của Trump.[17]

Trong nhiều thập kỷ, Quốc hội Hoa Kỳ lo ngại rằng sự dẫn đầu của Hoa Kỳ trong đổi mới đang bị đe dọa. Vào năm 1999, Đại diện Quận 13, Lynn Rivers (đảng Dân chủ) hỏi chuyên gia kinh tế Scott Stern rằng chính phủ Hoa Kỳ nên làm gì để giải quyết sự gia tăng trong chi tiêu R&D của Nhật Bản, Đức và những nước khác. Phản hồi của ông là: “Điều đầu tiên chúng ta nên làm là gửi cho họ một lá thư cảm ơn. Đầu tư sáng tạo không phải tình huống thua thắng. Những người tiêu dùng Hoa Kỳ sẽ được hưởng lợi từ nhiều đầu tư ở các quốc gia khác… Đó là cuộc đua mà chúng ta đều có thể thắng.”[18]

Ngoài việc đầu tư vào nghiên cứu, Trung Quốc còn có một lợi thế thứ hai: quy mô. Máy dự đoán cần dữ liệu và Trung Quốc có nhiều người để cung cấp dữ liệu đó hơn bất kỳ nơi nào trên thế giới. Họ có nhiều nhà máy để đào tạo robot hơn, nhiều người sử dụng điện thoại thông minh hơn để đào tạo các sản phẩm tiêu dùng và nhiều bệnh nhân hơn để đào tạo các ứng dụng về y tế.[19] Kai-Fu Lee, một chuyên gia AI Trung Quốc, người sáng lập phòng thí nghiệm nghiên cứu Bắc Kinh của Microsoft, và chủ tịch sáng lập Google Trung Quốc, nhận xét, “Hoa Kỳ và Canada có những nhà nghiên cứu AI giỏi nhất trên thế giới, nhưng Trung Quốc có hàng trăm người giỏi, và nhiều dữ liệu hơn… AI là lĩnh vực nơi mà bạn cần phát triển thuật toán và dữ liệu cùng nhau; một số lượng lớn dữ liệu tạo nên sự khác biệt lớn.”[20]

Truy cập dữ liệu là nguồn lợi thế thứ ba của Trung Quốc. Ví dụ, một trong những kỹ sư cao cấp nhất của Microsoft, Qi Lu, rời khỏi Hoa Kỳ để đến Trung Quốc, thấy đất nước này là nơi tốt nhất để phát triển AI. Ông nhận xét: “Không phải tất cả đều là về công nghệ. Đó là về cấu trúc của môi trường – văn hóa, chế độ chính sách. Đây là lý do vì sao AI cộng với Trung Quốc, với tôi, là một cơ hội thú vị. Nó chỉ là các nền văn hóa khác nhau, các chế độ chính sách khác nhau và một môi trường khác.”[21]

Đây chắc chắn là trường hợp đi theo hướng các tính năng như nhận diện khuôn mặt. Trung Quốc, trái ngược với Hoa Kỳ, duy trì một cơ sở dữ liệu hình ảnh khổng lồ cho việc nhận diện. Điều này cho phép các công ty khởi nghiệp như Face++ phát triển và cấp phép nhận diện khuôn mặt AI để khách hàng xác thực người lái xe với Didi, công ty lái xe lớn nhất ở Trung Quốc và họ cũng có thể chuyển tiền qua Alipay, ứng dụng thanh toán tiền được sử

dụng bởi hơn 120 triệu người ở Trung Quốc. Hệ thống này dựa hoàn toàn vào phân tích khuôn mặt để cho phép thanh toán. Hơn nữa, Baidu đang sử dụng AI nhận diện khuôn mặt để xác thực khách hàng thu mua vé máy bay và khách du lịch đến các điểm tham quan.[22]

Những yếu tố này có thể tạo ra một cuộc đua khi các quốc gia cạnh tranh để làm giãn những hạn chế về quyền riêng tư để cải thiện vị trí AI của họ. Tuy nhiên, công dân và người tiêu dùng coi trọng vấn đề quyền riêng tư. Có một sự đánh đổi cơ bản giữa sự xâm nhập với sự cá nhân hóa và tiềm năng cho sự không hài lòng ở khách hàng liên quan đến việc thu mua dữ liệu người dùng. Đồng thời, một lợi ích tiềm năng phát sinh từ khả năng cá nhân hóa tốt hơn của các dự đoán. Sự đánh đổi trở nên phức tạp hơn vì hiệu ứng “kẻ hưởng thụ miễn phí”[*]. Những người dùng muốn những sản phẩm tốt hơn được đào tạo thông qua dữ liệu cá nhân, nhưng họ thích những dữ liệu đó thu nhập từ những người khác; không phải họ.

[*] *Trong kinh tế học, "kẻ hưởng thụ miễn phí" chỉ những người thụ hưởng các lợi ích từ hàng hóa công cộng mà không chịu chi phí ít hơn so với lợi ích họ được hưởng (Nguồn: Wikipedia)*

Nhà khoa học máy tính Oren Etzioni lập luận rằng các hệ thống AI không nên “giữ lại hoặc tiết lộ những thông tin bảo mật mà không có sự chấp thuận rõ ràng từ nguồn của thông tin đó.“[23] Với việc Amazon Echo lắng nghe mọi cuộc nói chuyện trong nhà, bạn sẽ muốn có một vài sự kiểm soát. Điều này dường như rất hiển nhiên. Tuy nhiên, nó không đơn giản như vậy. Thông tin ngân hàng của bạn là bảo mật, nhưng vậy còn âm nhạc mà bạn nghe hoặc những chương trình truyền hình mà bạn xem? Thậm chí, bất kỳ khi nào bạn hỏi Echo một câu hỏi, nó có thể trả lời bằng một câu hỏi khác: “Bạn có đồng ý cho Amazon truy cập vào câu hỏi của bạn để tìm câu trả lời không?” Đọc qua tất cả những chính sách quyền riêng tư của tất cả các công ty thu thập dữ liệu có thể mất tới vài tuần.[24] Mỗi khi AI yêu cầu sự đồng ý để sử dụng dữ liệu của bạn, sản phẩm trở nên tồi tệ hơn. Nó làm gián đoạn trải nghiệm của người dùng. Nếu mọi người không cung cấp dữ liệu, thì AI không thể học được từ phản hồi, hạn chế khả năng tăng năng suất và gia tăng thu nhập của nó.

Một công nghệ mới nổi – blockchain – cung cấp một cách phân cấp cơ sở dữ

liệu và giảm chi phí xác minh dữ liệu. Công nghệ như vậy có thể kết hợp với AI để vượt qua những lo lắng về quyền riêng tư (và an ninh), đặc biệt khi chúng đã được sử dụng cho những giao dịch tài chính, lĩnh vực mà những vấn đề này rất quan trọng.[25] Ngay cả khi có đủ người cung cấp dữ liệu để AI có thể học, điều gì sẽ xảy ra nếu những người dùng đó không giống với mọi người? Giả sử chỉ có những người giàu có từ California và New York cung cấp dữ liệu cho máy dự đoán vậy sau đó, AI sẽ học cách phục vụ cho những cộng đồng này. Nếu mục đích của việc hạn chế thu thập dữ liệu cá nhân là để bảo vệ những người dễ bị tổn thương, thì sau đó nó sẽ tạo ra một vấn đề mới: những người dùng không được hưởng lợi từ những sản phẩm tốt hơn và sự giàu có hơn mà AI cung cấp.

Tận thế sắp đến rồi?

Liệu AI có phải là mối đe dọa đối với sự tồn tại của nhân loại? Ngoài việc đơn giản là liệu một AI trở nên bất hợp tác như Hal 9000 (trong *2001: A Space Odyssey*), điều khiến những người rất nghiêm túc và thông minh như Elon Musk, Bill Gates và Stephen Hawking thao thức là liệu chúng ta có sáng tạo ra một thứ gì đó giống như Skynet từ bộ phim *Terminator*. Họ lo sợ rằng một “siêu trí tuệ “ – thuật ngữ được đặt ra bởi nhà triết học đến từ Oxford, Nick Bostrom – sẽ nhanh chóng coi nhân loại như một mối đe dọa, một thứ gì đó ngứa mắt…[26] Hay nói cách khác, AI có thể là sự đổi mới công nghệ cuối cùng của chúng ta.[27]

Chúng tôi không ở vị trí có thể phân xử vấn đề này và chúng tôi thậm chí còn không thể cùng nhau đưa ra kết luận. Nhưng điều khiến chúng tôi ngạc nhiên là cuộc tranh luận này thực tế gắn liền với kinh tế tới mức nào: nó chính là sự cạnh tranh củng cố tất cả. Một siêu trí tuệ là một AI vượt trội hơn con người ngay cả ở những công việc đòi hỏi nhận thức cao nhất và có thể giải quyết mọi vấn đề một cách lý trí. Cụ thể, nó có thể phát minh và tự cải thiện. Trong khi tác giả khoa học viễn tưởng Vernor Vinge gọi điểm đó là “Điểm kì dị” và nhà nghiên cứu tương lai Ray Kurzweil đề xuất rằng con người không được chuẩn bị để nhìn thấy trước điều gì sẽ xảy ra ở thời điểm này bởi vì chúng ta theo định nghĩa không thông minh bằng cỗ máy, hóa ra các chuyên gia kinh tế lại được chuẩn bị khá tốt để nghĩ về điều này.

Trong nhiều năm, các chuyên gia kinh tế đã phải đối mặt với chỉ trích rằng

những tác nhân mà chúng tôi dựa vào để đưa ra giả thuyết đều là mô hình siêu thực và phi thực tế với hành vi con người. Cũng đúng, nhưng khi nói về siêu trí tuệ, điều đó có nghĩa là chúng tôi đã đi đúng hướng. Chúng tôi đã giả định trí tuệ siêu phàm trong phân tích của mình. Chúng tôi có được sự hiểu biết thông qua bằng chứng toán học, một tiêu chuẩn độc lập trí tuệ của sự thật. Quan điểm này khá hữu ích. Các chuyên gia kinh tế nói cho chúng ta biết rằng nếu một siêu trí tuệ muốn kiểm soát thế giới, nó sẽ cần tài nguyên. Vũ trụ có rất nhiều tài nguyên, nhưng ngay cả một siêu trí tuệ cũng phải tuân thủ quy luật của vật lý. Việc thu thập tài nguyên rất tốn kém. Do đó, kinh tế học cung cấp cái nhìn sâu sắc để hiểu một xã hội AI siêu trí tuệ sẽ phát triển như thế nào.

Đúng là nghiên cứu đang được tiến hành để máy dự đoán hoạt động trong bối cảnh rộng hơn, nhưng sự đột phá có khả năng mở đường cho trí tuệ nhân tạo vẫn còn chưa được khám phá. Trong một tài liệu chính sách được chuẩn bị bởi giám đốc điều hành của văn phòng tổng thống Hoa Kỳ, Ủy ban Công nghệ của Hội đồng Khoa học và Công nghệ Quốc gia (NSTC) nói rằng: "Sự đồng thuận hiện tại của cộng đồng chuyên gia tư nhân, mà Ủy ban của NSTC cũng tán thành, là sự phổ biến của AI sẽ không thể đạt được trong ít nhất vài thập kỷ tới. Ủy ban Công nghệ NSTC đánh giá là mối quan tâm lâu dài về siêu trí tuệ AI sẽ có ít ảnh hưởng lên chính sách hiện tại."[28]

Liệu đây có phải là kết thúc của thế giới mà chúng ta biết? Chưa chắc, nhưng nó là kết thúc của cuốn sách này. Các công ty đang khai thác AI ngay lúc này. Bằng cách áp dụng những nguyên lý kinh tế học đơn giản có khả năng củng cố việc dự đoán với chi phí thấp và những yếu tố bổ sung có giá trị cao, doanh nghiệp của bạn có thể đưa ra những lựa chọn và quyết định chiến lược tối ưu hóa ROI liên quan đến AI. Khi chúng ta chuyển từ máy dự đoán sang trí tuệ nhân loại thông thường hoặc thậm chí siêu trí tuệ, cho dù vào thời điểm nào đi chăng nữa, thì chúng ta sẽ ở một thời điểm AI khác. Đó là điều mà tất cả đều đồng ý. Khi sự kiện đó xảy ra, chúng tôi có thể tự tin dự đoán rằng kinh tế học sẽ không còn đơn giản như vậy nữa.

NHỮNG ĐIỂM CHÍNH

• Sự phát triển của AI đưa ra nhiều sự lựa chọn cho xã hội. Mỗi lựa chọn đại diện cho một sự đánh đổi. Ở giai đoạn này, trong khi công nghệ vẫn còn đang

chưa phát triển, có ba sự đánh đổi đặc biệt nổi bật ở cấp độ xã hội.

• Sự đánh đổi đầu tiên là năng suất so với sự phân phối. Nhiều người đề xuất rằng AI sẽ khiến chúng ta nghèo hơn hoặc thậm chí tệ hơn. Điều đó không đúng. Các chuyên gia kinh tế đều cho rằng tiến bộ công nghệ khiến chúng ta tốt hơn và nâng cao năng suất. Rõ ràng là AI sẽ nâng cao năng suất. Vấn đề không phải là tạo ra sự giàu có; mà là sự phân phối. AI có thể khiến vấn đề bất bình đẳng thu nhập tồi tệ hơn bởi hai lý do. Đầu tiên, bằng cách tiếp nhận thêm một số công việc cụ thể, AI có thể gia tăng sự cạnh tranh giữa con người trong những công việc còn lại, làm giảm mức lương và tiếp tục làm giảm phần trăm thu nhập. Thứ hai, máy dự đoán, cũng như các công nghệ liên quan đến máy tính khác, cần những kỹ năng cụ thể để vận hành, do đó AI có thể sẽ chỉ có lợi cho những người có tay nghề cao.

• Sự đánh đổi thứ hai là sự đổi mới so với sự cạnh tranh. Giống như phần lớn các công nghệ liên quan đến phần mềm, AI có nền kinh tế quy mô. Hơn nữa, các công cụ AI thường được đặc trưng bởi việc gia tăng lợi nhuận ở một mức độ nào đó: sự dự đoán chính xác hơn dẫn đến nhiều người sử dụng hơn, nhiều người sử dụng hơn tạo ra nhiều dữ liệu hơn, và nhiều dữ liệu hơn dẫn đến sự dự đoán chính xác hơn. Các doanh nghiệp có những động lực lớn hơn để xây dựng máy dự đoán nếu họ có quyền kiểm soát hơn, nhưng, điều đó có thể dẫn đến sự độc quyền. Sự đổi mới nhanh hơn có thể có lợi cho xã hội từ điểm nhìn ngắn hạn nhưng sẽ không tối ưu từ điểm nhìn xã hội hoặc dài hạn.

• Sự đánh đổi thứ ba là hiệu suất so với quyền riêng tư. AI sẽ hoạt động tốt hơn nếu có nhiều dữ liệu hơn. Cụ thể, chúng sẽ cá nhân hóa sự dự đoán tốt hơn nếu chúng có quyền truy cập tới nhiều dữ liệu cá nhân. Cung cấp dữ liệu cá nhân sẽ thường đi kèm với giảm quyền riêng tư. Một vài khu vực pháp lý, như ở châu Âu, đã chọn tạo ra một môi trường cung cấp cho công dân nhiều quyền riêng tư hơn. Điều này có thể có lợi cho công dân của họ và có thể tạo ra những điều kiện cho một thị trường năng động hơn cho thông tin cá nhân trong khi các cá nhân có thể dễ dàng quyết định liệu họ có muốn giao dịch, bán, hoặc quyên góp dữ liệu cá nhân. Mặc khác, điều này có thể tạo ra các sự xung đột trong những bối cảnh mà việc lựa chọn tham gia là tốn kém và gây bất lợi cho các công ty châu Âu cũng như công dân trong thị trường mà AI với quyền truy cập tới dữ liệu tốt hơn sẽ cạnh tranh hơn.

• Đối với cả ba đánh đổi, các khu vực pháp lý sẽ phải cân nhắc cả hai mặt của

giao dịch và thiết kế chính sách phù hợp nhất với chiến lược tổng thể và ý muốn của công dân.

Lời cảm ơn

Chúng tôi muốn bày tỏ lời cảm ơn tới những người đã đóng góp thời gian, ý tưởng và sự nhẫn nại cho cuốn sách này. Cụ thể, chúng tôi xin cảm ơn Abe Heifets của Atomwise, Liran Belanzon của BenchSci, Alex Shevchenko của Grammarly, Marc Ossip và Ben Edelman vì đã dành thời gian cho các buổi phỏng vấn với chúng tôi, và Kevin Bryan vì những lời nhận xét cho bản thảo. Ngoài ra, chúng tôi xin cảm ơn những người đồng nghiệp vì những buổi thảo luận và những phản hồi, bao gồm Nick Adams, Umair Akeel, Susan Athey, Naresh Bangia, Nick Beim, Dennis Bennie, James Bergstra, Dror Berman, Vincent Bérubé, Jim Bessen, Scott Bonham, Erik Brynjolfsson, Andy Burgess, Elizabeth Caley, Peter Carrescia, Iain Cockburn, Christian Catalini, James Cham, Nicolas Chapados, Tyson Clark, Paul Cubbon, Zavain Dar, Sally Daub, Dan Debow, Ron Dembo, Helene Desmarais, JP Dube, Candice Faktor, Haig Farris, Chen Fong, Ash Fontana, John Francis, April Franco, Suzanne Gildert, Anindya Ghose, Ron Glozman, Ben Goertzel, Shane Greenstein, Kanu Gulati, John Harris, Deepak Hegde, Rebecca Henderson, Geoff Hinton, Tim Hodgson, Michael Hyatt, Richard Hyatt, Ben Jones, Chad Jones, Steve Jurvetson, Satish Kanwar, Danny Kahneman, John Kelleher, Moe Kermani, Vinod Khosla, Karin Klein, Darrell Kopke, Johann Koss, Katya Kudashkina, Michael Kuhlmann, Tony Lacavera, Allen Lau, Eva Lau, Yann LeCun, Mara Lederman, Lisha Li, Ted Livingston, Jevon MacDonald, Rupam Mahmood, Chris Matys, Kristina McElheran, John McHale, Sanjog Misra, Matt Mitchell, Sanjay Mittal, Ash Munshi, Michael Murchison, Ken Nickerson, Olivia Norton, Alex Oettl, David Ossip, Barney Pell, Andrea Prat, Tomi Poutanen, Marzio Pozzuoli, Lally Rementilla, Geordie Rose, Maryanna Saenko, Russ Salakhutdinov, Reza Satchu, Michael Serbinis, Ashmeet Sidana, Micah Siegel, Dilip Soman, John Stackhouse, Scott Stern, Ted Sum, Rich Sutton, Steve Tadelis, Shahram Tafazoli, Graham Taylor, Florenta Teodoridis, Richard Titus, Dan Trefler, Catherine Tucker, William Tunstall-Pedoe, Stephan Uhrenbacher, Cliff van der Linden, Miguel Villas-Boas, Neil Wainwright, Boris Wertz, Dan Wilson, Peter Wittek, Alexander Wong, Shelley Zhuang và Shivon Zilis. Chúng tôi cũng muốn cảm ơn Carl Shapiro và Hal Varian vì cuốn sách Information Rules (tạm dịch: Tầm quan trọng của

thông tin) của họ đã truyền cảm hứng cho dự án của chúng tôi. Các nhân viên tại The Creative Destruction Lab và Rotman School rất tuyệt vời, đặc biệt là Steve Arenburg, Dawn Bloomfi eld, Rachel Harris, Jennifer Hildebrandt, Anne Hilton, Justyna Jonca, Aidan Kehoe, Khalid Kurji, Mary Lyne, Ken McGuffin, Shray Mehra, Daniel Mulet, Jennifer O'Hare, Gregory Ray, Amir Sariri, Sonia Sennik, Kristjan Sigurdson, Pearl Sullivan, Evelyn Thomasos và những nhân viên còn lại của đội ngũ Lab và Rotman. Chúng tôi gửi lời cảm ơn tới Trưởng khoa Tiff Macklem, vì sự hỗ trợ nhiệt tình cho công trình của chúng tôi về AI tại Creative Destruction Lab và trong suốt thời gian tại Rotman School. Xin cảm ơn ban lãnh đạo và nhân viên tại The Next 36 và The Next AI. Chúng tôi cũng muốn cảm ơn Walter Frick và Tim Sullivan vì khả năng chỉnh sửa tuyệt vời, cũng như đại diện của chúng tôi, Jim Levine.

Rất nhiều ý tưởng trong cuốn sách được dựa trên những nghiên cứu với sự giúp đỡ của Hội đồng Nghiên cứu Khoa học xã hội và Khoa học nhân văn Canada, Viện Vector, Viện Nghiên cứu cao cấp Canada dưới sự lãnh đạo của Alan Bernstein và Rebecca Finlay và Sloan Foundation, với sự giúp đỡ của Danny Goroff, dưới sự trợ cấp của Economics of Digitization, được quản lý bởi Shane Greenstein, Scott Stern, và Josh Lerner. Chúng tôi rất biết ơn sự giúp đỡ của họ. Chúng tôi cũng cảm ơn Jim Poterba vì sự giúp đỡ của ông cho buổi hội nghị của chúng tôi về các khía cạnh kinh tế của AI thông qua cục Nghiên cứu kinh tế quốc gia. Cuối cùng, chúng tôi gửi lời cảm ơn tới gia đình vì sự nhẫn nại và đóng góp trong suốt quá trình: Gina, Amelia, Andreas, Rachel, Anna, Sam, Ben, Natalie, Belanna, Ariel, Annika.

Chú thích

C

hương 2

1. Stephen Hawking, Stuart Russell, Max Tegmark, và Frank Wilczek, "Stephen Hawking: "Transcedence Looks at the Implications of Artificial Intelligence— But Are We Taking AI Seriously Enough?" The Independent, Ngày 1 Tháng 5, 2014, http://www.independent. co.uk/news/science/ stephen- hawking-transcendence- looks- at- the- implications- of- Artificial- intelligence- but-are-we-taking-9313474.html.

2. Paul Mozur, "Beijing Wants A.I. to Be Made in China by 2030," New York Times, Ngày 20 Tháng 7, 2017, https://www.nytimes.com/2017/07/20/business/ china-artificial-intelligence.html?mcubz=0&_r=0.

3. Steve Jurvetson, "Intelligence Inside," Medium , Ngày 9 Tháng 2016 ,8, https://medium.com/@DFJvc/ intelligence-inside54-dcad8c4a3e.

4. William D. Nordhaus, "Do Real- Output and Real-Wage Measures Capture Reality? The History of Lighting Suggests Not," Tổ chức nghiên cứu kinh tế Cowles, Đại học Yale, 1998, https://lucept.fi les. wordpress.com/2014/11/william-nordhaus-the-cost- of-light.pdf.

5. Đây là một phần của một xu hướng kéo dài về việc giảm thiểu chi phí chung cho máy tính. Xem thêm William D. Nordhaus, "Two Centuries of Productivity Growth in Computing," Journal of Economic History , vol. 67/1 (2007): 128–159.

6. Lovelace, trong Walter Isaacson, The Innovators: How a Group of Hackers, Geniuses, and Geeks Created the Digital Revolution (New York: Simon & Schuster, 2014), 27.

7. Như trên, trang 29.

8. Amazon đang nghiên cứu về những vấn đề bảo mật khả thi với kế hoạch đó. Vào năm 2017, họ cho ra mắt Amazon Key, một hệ thống cho phép người vận chuyển được mở cửa nhà bạn và giao hàng, tất cả đều được ghi lại bởi máy ghi hình để đảm bảo rằng mọi chuyện diễn ra ổn thỏa.

9. Thật thú vị là một vài công ty khởi nghiệp đã nghĩ theo cách này. Stitch Fix sử dụng máy tự học để dự đoán quần áo mà khách hàng sẽ muốn và gửi hàng đến cho họ. Khách hàng sẽ gửi lại quần áo mà họ không muốn. Vào 2017, Stitch Fix đã có một IPO thành công dựa trên mô hình này – có lẽ là công ty khởi nghiệp "ưu tiên AI" đầu tiên làm được điều này.

10. Xem thêm US Patent Number 8,615,473 B2 và Praveen Kopalle, "Why Amazon's Anticipatory Shipping is Pure Genius," Forbes , Ngày 28 Tháng 2014 ,1, https:// www.forbes.com/sites/onmarketing/28 /01 /2014/ why-amazons-anticipatory-shipping-is-pure-genius/2#a3284174605.

Chương 3

1. Như là một sự nhắc nhở về tầm quan trọng của việc diễn đạt đúng sự dự đoán, chúng tôi nhận ra rằng lời tiên tri ở Delphi dự đoán rằng một đế chế lớn sẽ bị phá hủy nếu bị tấn công. Emboldened, nhà vua tấn công Persia, đã rất ngạc nhiên khi đế chế Lydian của mình cũng bị phá hủy. Lời dự đoán về cơ bản đã đúng, nhưng bị hiểu nhầm.

2. "Mastercard Rolls Out Artificial Intelligence across Its Global Network," thông cáo báo chí của Mastercard, Ngày 30 Tháng 11, 2016, https://newsroom. mastercard.com/press-releases/ mastercard- rolls- out-Artificial-intelligence-across-its-global- network/.

3. Adam Geitgey, "Machine Learning Is Fun, Part 5: Language Translation with Deep Learning and the Magic of Sequences," Medium , Ngày 21 Tháng ,8 2016, https://medium.com/@ageitgey/ machine- learning- is- fun- part- 5-language- translation- with- deep- learning- and-the-magic-of-sequences-2ace0acca0aa.

4. Yiting Sun, "Why 500 Million People in China Are Talking to This AI," MIT Technology Review , Ngày 14 Tháng 9, 2017 https://www.technologyreview. com/s/608841/ why- -500 million- people-

in- china- are-talking-to-this-ai/.

5. Salvatore J. Stolfo, David W. Fan, Wenke Lee, và Andreas L. Prodromidis, "Credit Card Fraud Detection Using Meta- Learning: Issues and Initial Results," AAAI Technical Report , WS- 97-07, 1997, http://www.aaai.org/Papers/Workshops/1997/WS-97-07/WS97-07-015.pdf, với tỉ lệ sai dương khoảng 15% đến 20%. Một ví dụ khác là E. Aleskerov, B. Freisleben, và B. Rao, "CARDWATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection," Computational Intelligence for Financial Engineering, 1997, http://ieeexplore. ieee.org/stamp/stamp.jsp?arnumber=618940. Chú ý rằng những sự so sánh này không hoàn toàn tương đương, bởi vì họ sử dụng những cụm dữ liệu đào tạo khác nhau. Tuy nhiên quan điểm chung về sự chính xác vẫn đúng.

6. Abhinav Srivastava, Amlan Kundu, Shamik Sural, và Arun Majumdar, "Credit Card Fraud Detection Using Hidden Markov Model," IEEE Transactions on Dependable and Secure Computing 5, no. 1 (Tháng 1 – Tháng 3 2008): 37–48, http://ieeexplore. ieee.org/stamp/stamp.jsp?arnumber=4358713. Xem thêm Jarrod West và Maumita Bhattacharya, "Intelligent Financial Fraud Detection: A Comprehensive Review," Computers & Security 57 (2016): 47–66, http://www.sciencedirect.com/science/article/pii/S0167404815001261.

7. Andrej Karpathy, "What I Learned from Competing against a ConvNet on ImageNet," Andrej Karthy (blog), Ngày 2 Tháng 9, 2014, http://karpathy.github. io/2014/09/02/ what- i- learned- from- competing-against-a-convnet-on-imagenet/; ImageNet, Large Scale Visual Recognition Challenge 2016, http:// image-net.org/challenges/LSVRC/2016/results; Andrej Karpathy, LISVRC 2014, http://cs.stan ford. edu/people/karpathy/ilsvrc/.

8. Aaron Tilley, "China's Rise in the Global AI Race Emerges as It Takes Over the Final ImageNet Competition," Forbes , Ngày 31 Tháng 7, 2017 https:// www.forbes.com/sites/aarontilley/31/07/2017/china- ai-imagenet/#dafa182170a8.

9. Dave Gershgorn, "The Data That Transformed AI Research— and Possibly the World," Quartz , Ngày 26 Tháng 7, 2017

https://qz.com/1034972/ the-data- that-changed-the-direction-of- ai-research-and- possibly-the-world/.

10. Định nghĩa theo Oxford English Dictionary

Chương 4

1. J. McCarthy, Marvin L. Minsky, N. Rochester, và Claude E. Shannon, “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence,” Ngày 31 Tháng 8, 1955, http://www-formal.stanford. edu/jmc/history/dartmouth/dartmouth.html.

2. Jeff Hawkins và Sandra Blakeslee, On Intelligence (New York: Times Books, 2004), 89.

3. McCarthy và cộng sự, “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.”

4. Ian Hacking, The Taming of Chance (Cambridge, UK: Cambridge University Press, 1990).

Chương 5

1. Hal Varian, bài giảng “Beyond Big Data,” National Association of Business Economists, San Francisco, Ngày 10 Tháng 9, 2013.

2. Ngai- yin Chan và Chi- chung Choy, “Screening for Atrial Fibrillation in 13,122 Hong Kong Citizens with Smartphone Electrocardiogram,” BMJ 103, no.1 (Tháng 1 năm 2017), http://heart.bmj.com/ content/103/1/24; Sarah Buhr, “Apple’s Watch Can Detect an Abnormal Heart Rhythm with 97% Accuracy, UCSF Study Says,” Techcrunch, Ngày 11 Tháng 5, 2017 https://techcrunch.com/2017/05/11/ apples- watch-can- detect- an- abnormal-heart- rhythm- with-97-accuracy-ucsf-study-says/; Alive-Cor, “AliveCor and Mayo Clinic Announce Collaboration to Identify Hidden Health Signals in Humans,” Cision PR newswire, Ngày 24 Tháng 10, 2016, http://www.prnews wire.com/ news- releases/ alivecor- and- mayo- clinic- announce-collaboration- to-identify- hidden-health-signals-in- humans-300349847.html.

3. Buhr, "Apple's Watch Can Detect an Abnormal Heart Rhythm with 97% Accuracy, UCSF Study Says"; và Avesh Singh, "Applying Artificial Intelligence in Medicine: Our Early Results," Cardiogram (blog), Ngày 11 Tháng 5, https://blog.cardiogr.am/ applying- Artificial- intelligence- in-medicine-our- early-results-78bfe7605d32.

4. Chúng tôi không biết liệu Cardiogram có thành công hay không. Tuy nhiên chúng tôi tự tin rằng điện thoại thông minh và những cảm biến khác sẽ được sử dụng cho chẩn đoán y tế.

5. 6.000 là một con số tương đối nhỏ cho nghiên cứu này, đó là lý do chính vì sao nghiên cứu này được coi là "sơ khai". Dữ liệu đủ cho mục đích ban đầu của Cardiogram bởi vì nó là nghiên cứu sơ khai để cung cấp bằng chứng cho khái niệm. Không có mạng sống của ai bị nguy hiểm. Để cho kết quả được chính xác, nó sẽ cần nhiều dữ liệu hơn.

6. Dave Heiner, "Competition Authorities and Search," Microsoft Technet (blog), Ngày 26 Tháng 2, 2010, https://blogs.technet.microsoft.com/microsoft_on_ the_issues/2010/02/26/competition-authorities-and- search/. Google tranh luận rằng Bing đủ lớn để đạt được lợi nhuận quy mô trong tìm kiếm.

Chương 6

1. 60% thời gian bạn chọn X và đúng 60%, trong khi 40% bạn chọn O và chỉ đúng 40%. Trung bình, đây là 0.62 + 0.42 = 0.52.

2. Amost Tversky và Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," Science 185, no. 4157 (1974): 1124–1131, https://people.hss.caltech. edu/~camerer/Ec101/JudgementUncertainty.pdf.

3. Xem thêm Daniel Kahneman, Thinking, Fast and Slow (New York: Farrar, Strauss và Giroux, 2011); và Dan Ariely, Predictably Irrational (New York: HarperCollins, 2009).

4. Michael Lewis, Moneyball (New York: Norton, 2003).

5. Đương nhiên, trong khi Moneyball dựa vào việc sử dụng số liệu truyền

thống, không có gì ngạc nhiên khi các đội hiện đang tìm kiếm những phương pháp máy tự học để thực hiện chức năng, thu thập nhiều dữ liệu hơn trong quá trình. Xem thêm Takashi Sugimoto, “AI May Help Japan’s Baseball Champs Rewrite ‘Moneyball,’” Nikkei Asian Review , Ngày 2 Tháng 5, 2016 http://asia.nikkei.com/Business/ Companies/AI- may- help- Japan-s-baseball-champs- rewrite-Moneyball.

6. Jon Kleinberg, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, và Sendhil Mullainathan, “Human Decisions and Machine Predictions,” số 23180, National Bureau of Economic Research, 2017.

7. Nghiên cứu cũng chỉ ra rằng thuật toán có thể làm giảm sự phân biệt chủng tộc.

8. Mitchell Hoffman, Lisa B. Kahn, và Danielle Li, “Discretion in Hiring,” Số 21709, National Bureau of Economic Research, Tháng 11 năm 2015, chỉnh sửa Tháng 2 2016.

9. Donald Rumsfeld, điểm tin, Bộ Quốc phòng Hoa Kỳ, Ngày 12 Tháng 2, 2002, https://en.wikipedia.org/wiki/There_are_known_knowns.

10. Bertrand Rouet-Leduc và cộng sự, “Machine Learning Predicts Laboratory Earthquakes,” Đại học Cornell, 2017, http://arxiv.org/abs/1702.05774.

11. Dedre Gentner và Albert L. Stevens, Mental Models (New York: Psychology Press, 1983); Dedre Gentner, “Structure Mapping: A Theoretical Model for Analogy,” Cognitive Science 7 (1983): 15–170.

12. Ngay cả khi máy tốt hơn trong những trường hợp như vậy, quy luật của tính xác suất đồng nghĩa với việc trong những thử nghiệm nhỏ, sẽ luôn có sự không chắc chắn. Do đó, khi dữ liệu hiếm hoi, máy dự đoán sẽ không chính xác. Máy có thể cung cấp những dấu hiệu khi máy đưa ra những dự đoán không chính xác. Như chúng tôi thảo luận trong chương 8, điều này tạo ra một vai trò cho con người trong việc đánh giá nên phản ứng thế nào với những dự đoán không chính xác.

13. Nassim Nicholas Taleb, The Black Swan (New York: Random House,

2007).

14. Trong chương trình Foundation của Isaac Asimov, sự dự đoán trở nên đủ mạnh mẽ đến mức nó có thể nhìn trước sự phá hủy của đế chế Galactic và những sự đau đớn khác nhau của xã hội là trọng tâm của câu chuyện. Tuy nhiên, quan trọng với cốt truyện là những dự đoán có thể không nhìn trước được sự phát triển của những người "đột biến". Sự dự đoán không nhìn trước được những sự kiện bất ngờ.

15. Joel Waldfogel, "Copyright Protection, Technological Change, and the Quality of New Products: Evidence from Recorded Music since Napster," Journal of Law and Economics 55, no. 4 (2012): 715–740.

16. Donald Rubin, "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," Journal of Educational Psychology 66, no. 5 (1974):688–701; Jerzy Neyman, "Sur les applications de la theorie des probabilites aux experiences agricoles: Essai des principes," luận văn tiến sĩ, 1923, trích từ tái bản tiếng Anh, D. M. Dabrowska, và T. P. Speed dịch, Statistical Science 5 (1923): 463–472.

17. Garry Kasporov, Deep Thinking (New York: Perseus Books, 2017), 99–100.

18. Google Panda, Wikipedia , https://en.wikipedia.org/ wiki/Google_Panda, Truy cập Ngày 26 Tháng 7, 2017. Được mô tả trong webmaster của Google, "What's It Like to Fight Webspam at Google?" YouTube, Ngày 12 Tháng 2, 2014, https://www.youtube.com/ watch?v=rr-Cye_mFiQ.

19. Ví dụ, xuất bản có sửa chữa vào Tháng 9 năm 2016: Ashitha Nagesh, "Now You Can Finally Get Rid of All Those Instagram Spammers and Trolls," Metro , Ngày 13 Tháng 9, 2016 http://metro.co.uk/13/09/2016/ now-you- can- fi nally-get- rid- of- all- those-instagram- spammers-and-trolls6125645-/. Và rồi, một lần nữa, vào Tháng 6 năm 2017: Jonathan Vanian, "Instagram Turns to Artificial Intelligence to Fight Spam and Offensive Comments," Fortune , Ngày 29 Tháng 6, 2017 http://fortune.com/29/06/2017/instagram-artificial-intelligence-offensive-comments/. Thử thách sử dụng máy dự đoán với những yếu tố chiến lược là vấn đề với lịch sử dài. Năm 1976, chuyên gia kinh tế Robert Lucas đưa ra ý

kiến liên quan đến chính sách kinh tế vĩ mô về lạm phát và những chỉ số kinh tế khác. Nếu con người thay đổi hành vi tốt hơn sau khi chính sách thay đổi, họ sẽ làm vậy. Lucas nhấn mạnh rằng ngay cả khi việc làm có xu hướng tăng cao khi lạm phát cao, nếu một ngân hàng thay đổi chính sách của gia tăng lạm phát, mọi người vẫn sẽ mong chờ lạm phát đó và mối quan hệ sẽ bị phá hủy. Vậy nên, thay vì một chính sách dựa vào ngoại suy từ dữ liệu trong quá khứ, ông lập luận rằng chính sách sẽ phải dựa vào sự hiểu biết về những lý do thúc đẩy đằng sau hành vi con người. Điều này được biết đến là "Lucas Critique". Xem thêm Robert Lucas, "Econometric Policy Evaluation: A Critique," *Carnegie- Rochester Conference Series in Public Policy* 1, no. 1 (1976): 19–46, https://ideas.repec.org/a/eee/ crcspp/v1y1976ip19-46.html. Chuyên gia kinh tế Tim Harford mô tả điều này theo cách khác: Fort Knox chưa từng bị cướp. Nên chi tiêu bao nhiêu vào việc bảo vệ Fort Knox? Bởi vì nó chưa từng bị cướp, chi tiêu vào an ninh không dự đoán sự giảm các vụ cướp. Máy dự đoán có thể sẽ đề xuất không chi tiêu bất kỳ gì. Vậy sao phải chi tiêu tiền khi an ninh không làm giảm các vụ cướp? Tim Harford, The Undercover Economist Strikes Back: How to Run— or Ruin— an Economy (New York: Riverhead Books, 2014).

20. Dayong Wang và cộng sự, "Deep Learning for Identifying Metastatic Breast Cancer," Camelyon Grand Challenge, Ngày 18 Tháng 6, 2016, https:// arxiv.org/pdf/1606.05718.pdf.

21. Charles Babbage, On the Economy of Machinery and Manufactures (London:Charles Knight Pall Mall East, 1832), 162.

22. Daniel Paravisini và Antoinette Schoar, "The Incentive Effect of IT: Randomized Evidence from Credit Committees," số 19303, National Bureau of Economic Research, Tháng 8 năm 2013.

23. Sự phân chia lao động "thông qua trước" như vậy được thấy ở nhiều sự khai thác máy dự đoán. Washington Post có một AI trong nội bộ có thể phát hành 850 câu chuyện vào năm 2016, nhưng mỗi câu chuyện đều được kiểm tra lại bởi một ai đó trước khi được phát hành. Quá trình tương tự cũng được khai thác bởi ROSS Intelligence cho hàng ngàn văn bản chính thống và biến chúng thành một bản tin ngắn. Xem thêm Miranda Katz, "Welcome to the Era of the AI Coworker," Wired , Ngày 15 Tháng 2017 ,11 https://www.wired.com/story/ welcome- to- the-era- of-the-ai-coworker/.

Chương 7

1. Jody Rosen, "The Knowledge, London's Legendary Taxi- Driver Test, Puts Up a Fight in the Age of GPS," New York Times, Ngày 10 Tháng 11, 2014, https://www.nytimes.com/2014/11/10/t-magazine/london-taxi-test-knowledge.html?_r=0.

2. Để xem sách, xem thêm Joshua S. Gans, Core Economics for Managers (Australia: Cengage, 2005).

3. Để xem là vì sao:

Trung bình "Mang theo" = (3/4)(Không bị ướt khi có ô) + (1/4)(Không bị ướt khi có ô) = (3/4)8 + (1/4)8 = 8

Trung bình "Không mang theo" = (3/4)(Không bị ướt khi không có ô) + (1/4)(Bị ướt) = (3/4)10 + (1/4)0 = 7.5

Chương 8

1. Andrew McAfee và Erik Brynjolfsson, Machine, Platform, Crowd: Harnessing Our Digital Future (New York: Norton, 2017), 72.

2. Đây là ví dụ lấy từ Jean- Pierre Dubé và Sanjog Misra, "Scalable Price Targeting," báo cáo công việc, Trường Kinh doanh Booth thuộc Đại học Chicago, 2017, http:// conference.nber.org/confer//2017/SI2017/PRIT/Dube_Misra.pdf.

Chương 9

1. Daisuke Wakabayashi, "Meet the People Who Train the Robots (to Do Their Own Jobs)," New York Times, Ngày 28 Tháng 4, 2017, https://www.nytimes. com/2017/04/28/technology/ meet- the- people- who-train- the- robots-to-do-their-own-jobs.html?_r=1.

2. Như trên.

3. Ben Popper, "The Smart Bots Are Coming and This One Is Brilliant," The Verge, Ngày 7 Tháng 4, 2016,

https://www.theverge.com/2016/4/7/11380470/ amy- personal-digital-assistant-bot-ai-conversational.

4. Ellen Huet, “The Humans Hiding Behind the Chatbots,” Bloomberg, Ngày 18 Tháng 4, 2016, https:// www.bloomberg.com/news/articles/2016-04-18/ the- humans-hiding-behind-the-chatbots.

5. Wakabayashi, “Meet the People Who Train the Robots (to Do Their Own Jobs)”.

6. Marc Mangel và Francisco J. Samaniego, “Abraham Wald’s Work on Aircraft Survivability,” Journal of the American Statistical Association 79, no. 386 (1984): 259–267.

7. Bart J. Bronnenberg, Peter E. Rossi, và Naufel J. Vilcassim, “Structural Modeling and Policy Simulation,” *Journal of Marketing Research* 42, no. 1 (2005): 22–26, http://journals.ama.org/doi/ abs/10.1509/jmkr.42.1.22.56887.

8. Jean Pierre Dubé và cộng sự “Recent Advances in Structural Econometric Modeling,” Marketing Letters 16, no. 3–4 (2005): 209–224, https://link.springer.com/article/10.1007%2Fs11002-005-5886-0?LI=true.

Chương 10

1. “Robot Mailman Rolls on a Tight Schedule,” Popular Science , Tháng 10 năm 1976, https://books.google. ca/books?id=HwEAAAAAMBAJ&pg=PA76&lpg=PA76&dq=mailmobile+robot&source=bl&ots=SHkkOiDv8K&sig=sYFXzvvZ8_GvOV8Gt30hoGrFhpk&hl=en&sa=X&ei=B3kLVYr7N8meNoLsg_AD&redir_esc=y#v=onepage&q=mailmobile%20robot&f=false.

2. George Stigler, thông qua lời kể tới các tác giả bởi Nathan Rosenberg vào năm 1991.

3. Nobel citation: “Studies of Decision Making Lead to Prize in Economics,” Viện Hàn lâm Khoa học Hoàng gia Thụy Điển, thông cáo báo chí, Ngày 16 Tháng 10, 1978, https://www.nobelprize.org/nobel_ prizes/economic-

sciences/laureates/1978/press. html. Turing award citation: Herbert Alexander Simon, A.M. Turing Award, 1975, http://amturing. acm.org/award_winners/simon_1031467.cfm. Xem thêm Herbert A. Simon, “Rationality as Process and as Product of Thought,” American Economic Review 68, no. 2 (1978): 1–16; Allen Nevell và Herbert A. Simon, “Computer Science as Empirical Inquiry,” Communications of the ACM 19, no. 3 (1976): 120.

4. Frederick Jelinek trong Roger K. Moore, “Results from a Survey of Attendees at ASRU 1997 and 2003,” INTERSPEECH-2005, Lisbon, Ngày 4–8 Tháng 9, 2005.

Chương 11

1. Jmdavis, “Autopilot worked for me today and saved an accident,” Tesla Motors Club (blog), Ngày 12 Tháng 12, 2016, https://teslamotorsclub.com/tmc/threads/ autopilot- worked- for- me-today-and-saved-an- accident.82268/.

2. Một vài tuần sau, một máy ghi hình của một tài xe khác bắt được hệ thống hoạt động: Fred Lambert, “Tesla Autopilot’s New Radar Technology Predicts an Accident Caught on Dashcamera a Second Later,” Electrek, Ngày 27 Tháng 12, 2016, https://electrek. co/2016/12/27/ tesla- autopilot-radar-technology- predict-accident-dashcam/.

3. NHTSA, “U.S. DOT and IIHS Announce Historic Commitment of 20 Automakers to Make Automatic Emergency Braking Standard on New Vehicles,” Ngày 17 Tháng 3, 2016, https://www.nhtsa.gov/ press- releases/ us- dot- and- iihs- announce-historic- commitment-20-automakers-make-automatic- emergency.

4. Kathryn Diss, “Driverless Trucks Move All Iron Ore at Rio Tinto’s Pilbara Mines, in World First,” ABC News , Ngày 18 Tháng 10, 2015 http://www.abc.net. au/news/18-10-2015/ rio- tinto-opens-worlds-fi rst-automated-mine/6863814.

5. Tim Simonite, “Mining 24 Hours a Day with Robots,” MIT Technology Review, Ngày 28 Tháng 12, 2016, h tt

ps://www.technologyreview.com/s/603170/ mining- 24-hours-a-day-with-robots/.

6. Samantha Murphy Kelly, “Stunning Underwater Olympics Shots Are Now Taken by Robots,” CNN , Ngày 9 Tháng 8, 2016, http://money.cnn.com/2016/08/08/technology/olympics-underwater- robots-getty/.

7. Hoang Le, Andrew Kang, và Yisong Yue, “Smooth Imitation Learning for Online Sequence Prediction,” International Conference on Machine Learning, Ngày 19 Tháng 6, 2016, https://www.disneyresearch.com/publication/smooth-imitation-learning/.

8. Luật bao gồm (1) Robot không được làm tổn thương con người hoặc, khi không hành động, cho phép con người tới làm hại; (2) Robot phải tuân thủ các yêu cầu được đưa ra bởi con người trừ khi những yêu cầu đó trái ngược với điều thứ 1; (3) Robot phải bảo vệ sự tồn tại của nó trừ khi sự bảo vệ đó trái ngược với điều thứ 1 hoặc điều thứ 2. Xem thêm Isaac Asimov, “Runaround,” I, Robot (The Isaac Asimov Collection ed.) (New York: Doubleday, 1950), 40.

9. Department of Defense Directive 3000.09: Autonomy in Weapon Systems, Ngày 21 Tháng 11, 2012, https:// www.hsdl.org/?abstract&did=726163.

10. Ví dụ, có rất nhiều điều kiện cho phép những sự lựa chọn khác khi bị giới hạn thời gian khi chiến đấu. Mark Guburd, “Why Should We Ban Autonomous Weapons? To Survive,” IEEE Spectrum , Ngày 1 Tháng 6, 2016 http://spectrum.ieee.org/automaton/robotics/ military- robots/ why-should- we-ban- autonomous-weapons-to-survive.

Chương 12

1. Robert Solow, “We’d Better Watch Out,” New York Times Book Review, Ngày 12 Tháng 7, 1987, 36.

2. Michael Hammer, “Reengineering Work: Don’t Automate, Obliterate,” Harvard Business Review, Tháng 7–8, 1990, https://hbr.org/1990/07/reengineering-work-dont-automate-obliterate.

3. Art Kleiner, “Revisiting Reengineering,” Strategy + Business , Tháng 7 năm 2000, https://www.strategy- business.com/article/19570?gko=e05ea.

4. Nanette Byrnes, “As Goldman Embraces Automation, Even the Masters of the Universe Are Threatened,” MIT Technology Review, Ngày 7 Tháng 2, 2017, https:// www.technologyreview.com/s/603431/ as- goldman- embraces-automation- even-the- masters-of-the- universe-are-threatened/.

5. “Google Has More Than 1,000 Artificial Intelligence Projects in the Works,” The Week , Ngày 18 Tháng 10, 2016 http://theweek.com/speedreads/654463/ google- more-than -1000 -Artificial-intelligence-projects- works.

6. Scott Forstall, trong “How the iPhone Was Born,” video của Wall Street Journal, Ngày 25 Tháng 6, 2017, http:// www.wsj.com/video/ how- the-iphone- was- born- inside- stories-of- missteps- and-triumphs/302CFE23-392D-4020-B1BD-B4B9CEF7D9A8.html.

Chương 13

1. Steve Jobs trong Memory and Imagination: New Pathways to the Library of Congress, Michael Lawrence Films, 2006, https://www.youtube.com/watch?v=ob_ GX50Za6c.

Chương 14

1. Steven Levy, “A Spreadsheet Way of Knowledge,” Wired , Ngày 24 Tháng 10, 2014 https://backchannel. com/ a-spreadsheet-way-of-knowledge-8de60af7146e.

2. Nick Statt, “The Next Big Leap in AI Could Come from Warehouse Robots,” The Verge, Ngày 1 Tháng 6, 2017, https://www.theverge.com/2017/6/1/15703146/ kindred-orb-robot-ai-startup-w arehouse- automation.

3. L. B. Lusted, “Logical Analysis in Roentgen Diagnosis,” Radiology 74 (1960): 178–193.

4. Siddhartha Mukherjee, “A.I. versus M.D.,” New Yorker, Ngày 3 Tháng 4, 2017, http://www.newyorker.com/ magazine/2017/04/03/ai-versus-md.

5. S. Jha và E. J. Topol, “Adapting to Artificial Intelligence: Radiologists and Pathologists as Information Specialists,” Journal of the American Medical Association 316, no. 22 (2016): 2353–2354.

6. Nhiều ý tưởng liên quan đến cuộc thảo luận của Frank Levy trong “Computers and the Supply of Radiology Services,” Journal of the American College of Radiology 5, no. 10 (2008): 1067–1072.

7. Xem thêm Verdict Hospital (http://www. hospitalmanagement.net/features/feature51500/) để xem bài phỏng vấn Chủ tịch American College of Radiology năm 2009. Hoặc, để tham khảo nguồn học thuật, xem thêm Leonard Berlin, “The Radiologist: Doctor’s Doctor or Patient’s Doctor,” American Journal of Roentgenology 128, no. 4 (1977), http://www. ajronline.org/doi/pdf/10.2214/ajr.128.4.702.

8. Levy, “Computers and the Supply of Radiology Services.”

9. Jha và Topol, “Adapting to Artificial Intelligence”; S. Jha, “Will Computers Replace Radiologists?” Medscape 30 (Tháng 12 năm 2016), http://www. medscape.com/viewarticle/863127#vp_1.

10. Carl Benedikt Frey và Michael A. Osborne, “The Future of Employment: How Susceptible Are Jobs to Computerisation?” Trường Oxford Martin, Đại học Oxford, Tháng 9 năm 2013, http://www. oxfordmartin.ox.ac.uk/downloads/academic/The_ Future_of_Employment.pdf.

11. Những nhà sản xuất xe tải đang tích hợp tính năng trong những phương tiện mới của họ. Volvo đã khai thác điều này trong nhiều bài kiểm tra, và loại xe bán tải mới của Telsa có những khả năng này ngay từ đầu.

Chương 15

1. “How Germany’s Otto Uses Artificial Intelligence,” The Economist, Ngày 12 Tháng 4, 2017, https://www. economist.com/news/business/21720675-

firm- using-algorithm- designed- cern-laboratory-how- germanys-otto-uses.

2. Zvi Griliches, “Hybrid Corn and the Economics of Innovation,” Science 29 (Tháng 7 năm 1960): 275–280.

3. Bryce Ryan và N. Gross, “The Diffusion of Hybrid Seed Corn,” Rural Sociology 8 (1943): 15–24; Bryce Ryan và N. Gross, “Acceptance and Diffusion of Hybrid Corn Seed in Two Iowa Communities,” Iowa Agriculture Experiment Station Research Bulletin, no. 372 (Tháng 1 năm 1950).

4. Kelly Gonsalves, “Google Has More Than 1,000 Artificial Intelligence Projects in the Works,” The Week, Ngày 18 Tháng 10, 2016, http://theweek. com/speedreads/654463/ google- more- than-1000- Artificial-intelligence-projects-works.

5. Một cuộc tranh cãi phong phú, mang tính giải trí và vô ích về liệu những chuyên gia phân tích thực nghiệm bóng chày giỏi hơn hay tốt hơn huấn luyện viên. Như Nate Silver nhấn mạnh, cả Moneyball và huấn luyện viên đều có những vai trò quan trọng. Nate Silver, The Signal and the Noise (New York: Penguin Books, 2015), chương 3.

6. Bạn có thể bắt gặp và nói rằng, để phát triển, máy dự đoán cần có dữ liệu từ quá khứ? Đây là một vấn đề nhạy cảm. Sự dự đoán hoạt động tốt nhất khi thêm dữ liệu mới không thay đổi thuật toán quá nhiều – sự ổn định là kết quả của sự thực hiện thống kê tốt. Điều đó có nghĩa là khi bạn sử dụng dữ liệu phản hồi để nâng cao thuật toán, nó có giá trị chính xác nhất khi điều được dự đoán tự phát triển. Vậy nếu nhu cầu cho sữa chua đột nhiên thay đổi với nhân khẩu học hoặc những yếu tố khác, dữ liệu mới sẽ giúp bạn thay đổi thuật toán. Tuy nhiên, nó chính xác khi những thay đổi đồng nghĩa với việc dữ liệu cũ ít hữu ích hơn cho sự dự đoán.

7. Daniel Ren, “Tencent Joins the Fray with Baidu in Providing Artificial Intelligence Applications for Self-Driving Cars,” South China Morning Post , Ngày 27 Tháng 8, 2107 http://www.scmp.com/business/ companies/article/2108489/tencent-forms-alliance- push-ai-applications-self-driving.

8. Ren, “Tencent Joins the Fray with Baidu in Providing Artificial Intelligence Applications for Self- Driving Cars.”

Chương 16

1. Lý thuyết của sự thích nghi và động lực được mô tả ở đây trích từ Steven Tadelis, “Complexity, Flexibility, and the Make- or- Buy Decision,” American Economic Review 92, no. 2 (Tháng 5 năm 2002): 433–437.

2. Silke Januszewski Forbes và Mara Lederman, “Adaptation and Vertical Integration in the Airline Industry,” American Economic Review 99, no. 5 (December 2009): 1831–1849.

3. Sharon Novak và Scott Stern, “How Does Outsourcing Affect Performance Dynamics? Evidence from the Automobile Industry,” Management Science 54, no.12 (Tháng 12 năm 2008): 1963–1979.

4. Jim Bessen, Learning by Doing (New Haven, CT: Yale University Press, 2106).

5. Vào 2016, Wells Fargo đối mặt với những cáo buộc gian lận lớn bởi những hành động của những người quản lý tài khoản khi mở những tài khoản tốn kém cho khách hàng và bắt họ trả phí.

6. Cuộc thảo luận này dựa trên Dirk Bergemann và Alessandro Bonatti, “Selling Cookies,” American Economic Journal: Microeconomics 7, no. 2 (2015): 259–294.

7. Một ví dụ là dịch vụ tư vấn Mastercard Advisors, sử dụng số lượng dữ liệu lớn của Mastercard để cung cấp một loạt sự dự đoán, từ gian lận khách hàng đến tỉ lệ giữ. Xem thêm http://www.mastercardadvisors. com/consulting.html.

Chương 17

1. Như nói với Steven Levy. Xem thêm Will Smith, “Stop Calling Google Cardboard’s 360-Degree Videos ‘VR,’” Wired , Ngày 16 Tháng 11, 2105 https://www. wired.com/-360/11/2015video-isnt-virtual-reality/.

2. Jessir Hempel, “Inside Microsoft’s AI Comeback,” Wired , Ngày 21 Tháng 6, 2107 https://www.wired. com/story/inside-microsofts-ai-comeback/.

3. “What Does It Mean for Google to Become an ‘ AI-First’ (Quoting Sundar) Company?” Quora, Tháng 4 năm 2016, https://www.quora.com/ What-does- it- mean-for- Google- to-become-an-AI-fi rst-company.

4. Clayton M. Christensen, The Innovator’s Dilemma (Boston: Harvard Business Review Press, 2016).

5. Để xem thêm về những tình huống nan giải gián đoạn này, xem thêm Joshua S. Gans, The Disruption Dilemma (Cambridge, MA: MIT Press, 2016).

6. Nathan Rosenberg, “Learning by Using: Inside the Black Box: Technology and Economics,” bài luận, Đại học Illinois tại Urbana–Champaign, 1982, 120–140.

7. Trong trường hợp trò chơi điện tử, bởi vì mục tiêu (điểm tuyệt đối) có liên quan gần gũi đến sự dự đoán (liệu bước đi này khiến điểm tăng hay giảm?), quá trình tự động không cần sự đánh giá riêng. Sự đánh giá là sự nhận ra đơn giản rằng mục tiêu là ghi được nhiều điểm nhất. Dạy cho máy cách chơi một trò chơi như Minecraft hoặc một trò chơi sưu tầm như Pokemon Go có thể đòi hỏi sự đánh giá nhiều hơn, bởi vì nhiều người khác nhau tận hưởng những khía cạnh khác nhau của trò chơi. Nó không rõ ràng rằng mục tiêu nên phải như thế nào.

8. Chesley “Sully” Sullenberger quoted in Katy Couric, “Capt. Sully Worried about Airline Industry,” CBS News , Ngày 10 Tháng 2, 2009; https://www.cbsnews. com/news/capt-sully-worried-about-airline- industry/.

9. Mark Harris, “Tesla Drivers Are Paying Big Bucks to Test Flawed Self-Driving Software,” Wired , Ngày 4 Tháng 3, 2107 https://backchannel.com/ tesla- drivers-are- guinea- pigs- for-fl awed-self-driving- software-c2cc80b483a#.s0u7lsv4f.

10. Nikolai Yakovenko, “GANS Will Change the World,” Medium , Ngày 3 Tháng 1, 2017, https:// medium.com/@Moscow25/ gans-will-change-the-

world-7ed6ae8515ca; Sebastian Anthony, “Google Teaches ‘AIs’ to Invent Their Own Crypto and Avoid Eavesdropping,” Ars Technica, Ngày 28 Tháng 10, 2016, https://arstechnica.com/information- technology/2016/10/google-ai-neural-network- cryptography/.

11. Apple, “Privacy,” https://www.apple.com/ca/privacy/.

12. Như trên.

13. Cuộc cá cược trở nên khả thi nhờ công nghệ phân tích dữ bảo mật tân tiến, đặc biệt là nhờ phát minh về phân loại các chính sách bảo mật của Cynthia Dwork: Cynthia Dwork, “Differential Privacy: A Survey of Results,” M. Agrawal, D. Du, Z. Duan, và A. Li (eds), Theory and Applications of Models of Computation. TAMC 2008. Lecture Notes in Computer Science , vol 4978 (Berlin: Springer, 2008), https://doi. org/10.1007/978-3-540-79228-4_1.

14. William Langewiesche, “The Human Factor,” Vanity Fair, Tháng 10 2014, http://www.vanityfair.com/ news/business/2014/10/air-france-flight-447-crash.

15. Tim Harford, “How Computers Are Setting Us Up for Disaster,” The Guardian, Ngày 11 Tháng 10, 2016, https:// www.theguardian.com/technology/2016/oct/11/ crash- how- computers-are-setting-us-up-disaster.

Chương 18

1. L. Sweeney, “Discrimination in Online Ad Delivery,” Communications of the ACM 56, no. 5 (2013): 44–54, https://dataprivacylab.org/projects/onlineads/.

2. Như trên.

3. “Racism Is Poisoning Online Ad Delivery, Says Harvard Professor,” MIT Technology Review, Ngày 4 Tháng 2, 2013, https://www.technologyreview. com/s/510646/ racism- is- poisoning- online-ad- delivery-says-harvard-professor/.

4. Anja Lambrecht và Catherine Tucker, "Algorithmic Bias? An Empirical Study into Apparent Gender- Based Discrimination in the Display of STEM Career Ads" (được trình bày tại NBER Summer Institute, Tháng 7 năm 2017).

5. Diane Cardwell và Libby Nelson, "The Fire Dept. Tests That Were Found to Discriminate," New York Times, Ngày 23 Tháng 7, 2009, https://cityroom.blogs. nytimes.com/2009/07/23/ the- fi re- dept- tests-that-were-found-to-discriminate/?mcubz=0&_r=0; US v. City of New York (FDNY), https://www.justice.gov/ archives/crt-fdny/overview.

6. Paul Voosen, "How AI Detectives Are Cracking Open the Black Box of Deep Learning," Science, Ngày 6 Tháng 7, 2017, http://www.sciencemag.org/news/2017/07/ how-ai- detectives- are- cracking-open-black-box-deep-learning.

7. T. Blake, C. Nosko, và S. Tadelis, "Consumer Heterogeneity and Paid Search Effectiveness: A Large-Scale Field Experiment," Econometrica 83 (2015): 155–174.

8. Hossein Hosseini, Baicen Xiao và Radha Poovendran, "Deceiving Google's Cloud Video Intelligence API Built for Summarizing Videos" (báo cáo tại Hội thảo CVPR, Ngày 31 Tháng 3, 2017), https://arxiv.org/ pdf/1703.09793.pdf; đồng thời hãy tìm đọc "Artificial Intelligence Used by Google to Scan Videos Could Easily Be Tricked by a Picture of Noodles," Quartz, Ngày 4 Tháng 4, 2017, https://qz.com/948870/ the- ai- used- by-google- to- scan- videos- could- easily- be- tricked-by-a-picture-of-noodles/.

9. Xem thêm, ví dụ, hàng ngàn trích dẫn tới C. S. Elton, The Ecology of Invasions by Animals and Plants (New York: John Wiley, 1958).

10. Dựa trên những cuộc thảo luận với Chủ tịch khoa của Đại học Waterloo Pearl Sullivan, giáo sư Alexander Wong, và những giáo sư khác của Waterloo vào Ngày 20 Tháng 11, 2016.

11. Có lợi ích thứ 4 cho việc dự đoán: đôi khi cần thiết cho những mục đích thực tiễn. Ví dụ, Google Glass cần quyết định liệu một chuyển động mí mắt là nháy mắt (không cố ý) hoặc nháy mắt (cố ý), với việc nháy mắt cố ý là một

cách để máy có thể bị kiểm soát. Bởi vì tốc độ mà sự quyết định cần được đưa ra, gửi dữ liệu tới đám mây và chờ câu trả lời là điều không thể. Máy dự đoán cần có trong thiết bị.

12. Ryan Singel, “Google Catches Bing Copying; Microsoft Says ‘So What?’”, Wired , Ngày 1 Tháng 2011 ,2, https:// www.wired.com/02/2011/bing-copies-google/.

13. Xem thêm Shane Greenstein cho một buổi thảo luận vì sao nó lại không được chấp nhận; “Bing Imitates Google: Their Conduct Crosses a Line,” Virulent Word of Mouse (blog), Ngày 2 Tháng 2, 2011, https:// virulentwordofmouse.wordpress.com/2011/02/02/ bing-imitates- google-their-conduct-crosses-a-line/; and Ben Edelman for a coun-terpoint, “In Accusing Microsoft, Google Doth Protest Too Much,” hbr.org, Ngày 3 Tháng 2, 2011, https://hbr.org/2011/02/in- accusing-microsoft-google.html.

14. Thật thú vị khi thấy nỗ lực của Google để kiểm soát máy tự học của Microsoft không thực sự hiệu quả. Trong số 100 thí nghiệm họ thực hiện, chỉ có 7 đến 9 thí nghiệm thực sự xuất hiện trên kết quả tìm kiếm của Bing. Xem thêm Joshua Gans, “The Consequences of Hiybbprqag’ing,” Digitopoly , Ngày 8 Tháng 2, 2011; https://digitopoly.org/2011/02/08/the-consequences-of-hiybbprqaging/.

15. Florian Tramèr, Fan Zhang, Ari Juels, Michael K. Reiter và Thomas Ristenpart, “Stealing Machine Learning Models via Prediction APIs” (báo cáo tại buổi điều trần của Hội nghị An ninh USENIX lần thứ 25, Austin, TX, Ngày 10–12 Tháng 8, 2016), https://regmedia.co.uk/2016/09/30/sec16_paper_tramer.pdf.

16. James Vincent, “Twitter Taught Microsoft’s AI Chatbot to Be a Racist Asshole in Less Than a Day,” The Verge , Ngày 24 Tháng 2016 ,3, https://www. theverge.com/11297050 /24 /3 /2016/tay-microsoft- chatbot-racist.

17. Rob Price, “Microsoft Is Deleting Its Chatbot’s Incredibly Racist Tweets,” Business Insider, Ngày 24 Tháng 3, 2016, http://www.businessinsider.com/ microsoft-deletes- racist- genocidal- tweets-from-ai- chatbot-tay-2016-3?r=UK&IR=T.

Chương 19

1. James Vincent, “Elon Musk Says We Need to Regulate AI Before It Becomes a Danger to Humanity,” The Verge , Ngày 17 Tháng 2017 ,7, https://www.theverge. com/15980954/17/7/2017/elon-musk-ai-regulation-existential-threat.

2. Chris Weller, “One of the Biggest VCs in Silicon Valley Is Launching an Experiment That Will Give 3000 People Free Money Until 2022,” Business Insider, Ngày 21 Tháng 9, 2017, http://www.businessinsider. com/y-combinator-basic-income-test-2017-9.

3. Stephen Hawking, “This Is the Most Dangerous Time for Our Planet,” The Guardian, Ngày 1 Tháng 12, 2016, https:// www.theguardian.com/commentisfree/2016/dec/01/ stephen-hawking-dangerous-time-planet-inequality.

4. “The Onrushing Wave,” The Economist, Ngày 18 Tháng 1, 2014, https://www.economist.com/ news/briefing/21594264-previous-technological- innovation-has-always-delivered-more-long-run- employment-not-less.

5. Để biết thêm, xem thêm John Markoff, Machines of Loving Grace: The Quest for Common Ground Between Humans and Robots (New York: Harper Collins, 2015); Martin Ford, Rise of the Robots: Technology and the Threat of a Jobless Future (New York: Basic Books, 2016); và Ryan Avent, The Wealth of Humans: Work, Power, and Status in the Twenty- First Century (London: St. Martin’s Press, 2016).

6. Jason Furman, “Is This Time Different? The Opportunities and Challenges of AI,” https:// obamawhitehouse.archives.gov/sites/default/fi les/ page/fi les/20160707 _cea_ai_furman.pdf.

7. Claudia Dale Goldin và Lawrence F. Katz, The Race between Education and Technology (Cambridge, MA: Harvard University Press, 2009), 90.

8. Lesley Chiou và Catherine Tucker, “Search Engines and Data Retention:Implications for Privacy and Antitrust,” báo cáo no. 23815, Văn

phòng Nghiên cứu Kinh tế Quốc gia, http://www.nber.org/papers/ w23815.

9. Google AdWords, “Reach more customers with broad match,” 2008.

10. Để xem những bài nhận xét về sự chống độc quyền và những ảnh hưởng khác xung quanh thuật toán, dữ liệu, và AI, xem thêm Ariel Ezrachi và Maurice Stucke, Virtual Competition: The Promise and Perils of the Algorithm- Driven Economy (Cambridge, MA: Harvard University Press, 2016). Để xem thêm cách nhìn mà liệu thuật toán sẽ được tập trung vào một thuật toán duy nhất, xem thêm Pedro Domingos, The Master Algorithm (New York: Basic Books, 2015). Cuối cùng, Steve Lohr cung cấp một cái nhìn tổng quan về cách các doanh nghiệp ưu tiên đầu tư vào dữ liệu cho lợi thế chiến lược; xem thêm Steve Lohr, Dataism (New York: Harper Business, 2015).

11. James Vincent, “Putin Says the Nation That Leads in AI ‘Will Be the Ruler of the World,’” The Verge, Ngày 4 Tháng 9, 2017, https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world.

12. Các báo cáo bao gồm: (1) Jason Furman, “Is This Time Different? The Opportunities and Challenges of Artificial Intelligence” (remarks at AI Now, Đại học New York, Ngày 7 Tháng 7, 2016), https:// obamawhitehouse.archives.gov/sites/default/ fi les/page/ fi les/20160707_cea_ai_furman.pdf; (2) Văn phòng điều hành của phủ tổng thống Hoa Kỳ, “Artificial Intelligence, Automation, and the Economy,” Tháng 10 năm 2016, https:// obamawhitehouse.archives.gov/sites/whitehouse. gov/ fi les/documents/Arti fi cial-Intelligence- Automation-Economy.PDF; (3) Văn phòng điều hành của phủ tổng thống Hoa Kỳ, Hội đồng Khoa học và Công nghệ Quốc gia, Uỷ ban Khoa học, “Preparing for the Future of Artificial Intelligence,” Tháng 10 năm 2016, https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/mic (4) Hội đồng Khoa học và Công nghệ Quốc gia, Viện nghiên cứu Mạng xã hội và Phát triển công nghệ thông tin, “The National Artificial Intelligence Research and Development Strategic Plan,” Tháng 10 năm 2016, https://obamawhitehouse.archives.gov/sites/default/ fi les/whitehouse_ fi les/microsites/ostp/NSTC/ national_ai_rd_strategic_plan.pdf.

13. Dan Trefler và Avi Goldfarb, “AI and Trade,” Ajay Agrawal, Joshua

Gans, vàAvi Goldfarb, trong cuốn sách sắp xuất bản Economics of AI.

14. Paul Mozur, “Beijing Wants AI to Be Made in China by 2030,” New York Times, Ngày 20 Tháng 7, 2017, https://www.nytimes.com/2017/07/20/business/ china-artificial-intelligence.html?_r=0.

15. “Why China’s AI Push Is Worrying,” The Economist, Ngày 27 Tháng 2017 ,7, https://www.economist. com/news/leaders/-21725561 state-controlled- corporations- are-developing- powerful- Artificial- intelligence-why-chinas-ai-push?frsc=dg7%Ce.

16. Paul Mozur, “Beijing Wants AI to Be Made in China by 2030,” New York Times, Ngày 20 Tháng 7, 2017, https://www.nytimes.com/2017/07/20/business/ china-artificial-intelligence.html?_r=0.

17. Như trên.

18. Ảnh 37 về tác động của nghiên cứu cơ bản trong đổi mới công nghệ và thịnh vượng quốc gia: Điều trần trước tiểu ban về nghiên cứu cơ bản của Ủy ban Khoa học, Hạ viện, Đại hội 106, phiên họp đầu tiên, Ngày 28 Tháng 9, 1999, 27.

19. “Why China’s AI Push Is Worrying.”

20. Will Knight, “China’s AI Awakening,” MIT Technology Review , Tháng 11 năm 2017.

21. Jessi Hempel, “How Baidu Will Win China’s AI Race— and Maybe the World’s,” Wired , Ngày 9 Tháng 2017 ,8, https://www.wired.com/story/ how- baidu- will- win- chinas-ai-raceand-maybe-the-worlds/.

22. Will Knight, “10 Breakthrough Technologies— 2017: Paying with Your Face,” MIT Technology Review, Tháng 3 – 4 2017, https://www.technologyreview. com/s/603494/10-breakthrough-technologies-2017- paying-with-your-face/.

23. Oren Etzioni, “How to Regulate Artificial Intelligence,” New York Times, Ngày 1 Tháng 9, 2017, https://www.nytimes.com/2017/09/01/opinion/artificial-intelligence-regulations-rules.html?_r=0.

24. Aleecia M. McDonald và Lorrie Faith Cranor, “The Cost of Reading Privacy Policies,” I/S 4, no. 3 (2008): 543–568, http://heinonline.org/HOL/Page?handle=hein.journals/isjlpsoc4&div=27&g_sent=1&casa_token=&collection=journals.

25. Christian Catalini và Joshua S. Gans, “Some Simple Economics of the Blockchain,” báo cáo no. 2874598, Trường Quản lý Rotman, Ngày 21 Tháng 9, 2017, và báo cáo No. 5191-16 của Viện nghiên cứu MIT Sloan, có tại https://ssrn.com/abstract=2874598.

26. Nick Bostrom, Superintelligence (Oxford, UK: Oxford University Press,2016).

27. Để biết thêm về những tranh luận xuất sắc về vấn đề này, hãy tìm đọc Max Tegmark, Life 3.0:Being Human in the Age of Artificial Intelligence (New York: Knopf, 2017).

28. “Prepare for the Future of Artificial Intelligence,” Giám đốc điều hành văn phòng, Hội đồng Khoa học và Công nghệ Quốc gia, Ủy ban Công nghệ, Tháng 10 năm 2016.